(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/065973 A2

(51) International Patent Classification?: A61K

(21) International Application Number: PCT/US02/34769

(22) International Filing Date: 28 October 2002 (28.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/348,434 26 October 2001 (26.10.2001) US

(71) Applicants *(for all designated States except US)*: ID BIOMEDICAL CORPORATION OF WASHINGTON [US/US]; 19204 North Creek Parkway, Suite 100, Bothell, WA 98011 (US). UNIVERSITY OF TENNESSEE RESEARCH CORPORATION [US/US]; 1534 White Avenue, Suite 403, Knoxville, TN 37996-1527 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: REDDISH, Mark, A. [US/US]; 18605 215th Way NE, Woodinville, WA 98072 (US). HU, Mary, Chaohong [CN/US]; 18209 76th Avenue West, Edmonds, WA 98026 (US). WALLS, Michael, A. [US/US]; 20040 61st Court Northeast, Kenmore, WA 98028-8533 (US). DALE, James, B. [US/US]; 72 Lombardy Road, Memphis, TN 38111 (US).

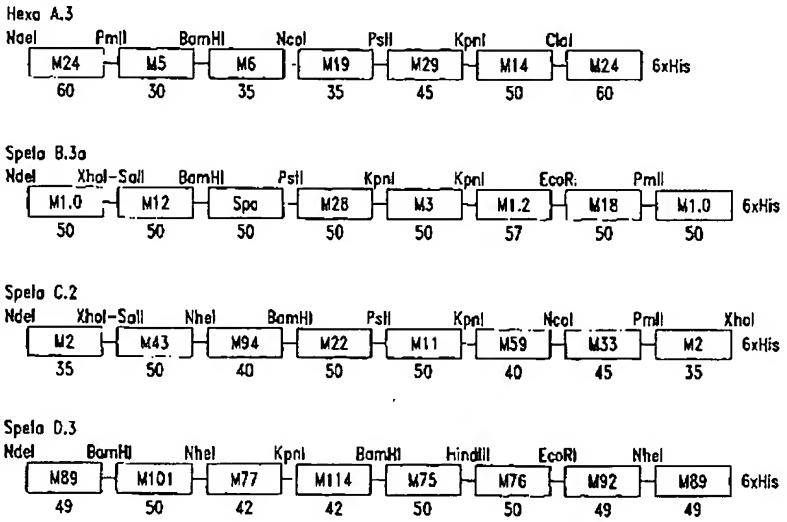
(74) Agents: PEPE, Jeffrey, C. et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,

[Continued on next page]

(54) Title: MULTIVALENT STREPTOCOCCAL VACCINE COMPOSITIONS AND METHODS FOR USE



(57) Abstract: Compositions and methods for making and using therapeutic formulations of multivalent hybrid polypeptides comprising immunogenic peptides of M proteins from various different serotypes of group A streptococci and antibodies thereto are provided. Also provided are nucleic acids encoding such hybrid polypeptides. The hybrid polypeptide formulations may be used, for example, in methods for treating or preventing a microbial infection and eliciting a protective immune response having broadly protective opsonic antibodies in the absence of tissue cross-reactive antibodies.

WO 03/065973 A2

TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MULTIVALENT STREPTOCOCCAL VACCINE COMPOSITIONS AND METHODS FOR USE

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Patent Application No. 60/348,434 filed October 26, 2001. This provisional application is incorporated herein by reference in its entirety.

STATEMENT OF GOVERNMENT INTEREST

This invention was made with government support under Grant No. AI-10085 awarded by the National Institutes of Health, and support from the Department of Veteran Affairs, Merit Review funds. The government may have certain rights in this invention.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to the prevention of infectious disease, and more specifically, to compositions comprising multivalent hybrid polypeptides having immunogenic M protein and Spa peptides capable of eliciting protective immunity.

DESCRIPTION OF THE RELATED ART

Group A streptococcal pharyngitis is one of the most common bacterial infections in school-age children. In addition to streptococcal pharyngitis, there may be associated nonsuppurative sequela, such as acute rheumatic fever (ARF). Although the incidence of ARF has declined in developed countries, the disease remains rampant in developing countries (Community prevention and control of cardiovascular diseases. Report of a WHO expert committee. World Health Organization. 1986:732). Another form of streptococcal infection is invasive, which afflicts thousands of children and adults each year, often resulting in death or significant morbidity (Stevens, *J. Infect. Dis.* 2:S366,

1999). Efforts to develop a vaccine that would prevent group A streptococcal infections have been ongoing for over eight decades (Dochez *et al.*, *J. Exp. Med. 30*:179, 1919; Lancefield, *J. Exp. Med. 47*:91, 1928).

Hence, a need exists for identifying and developing compositions therapeutically effective against streptococcal infections, particularly those compositions that can function as a vaccine and elicit protective immunity. Furthermore, there is a need for vaccine formulations that can be varied to protect against or treatment for infection by different streptococcal serotypes. The present invention meets such needs, and further provides other related advantages.

SUMMARY OF THE INVENTION

The present invention provides the discovery of therapeutic formulations of multivalent hybrid polypeptides, particularly a cocktail of hybrid polypeptides useful for eliciting a protective immune response having broadly protective opsonic antibodies in the absence of tissue cross-reactive antibodies. Hybrid polypeptides of the invention comprise at least six different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 30 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one serotype of group A streptococci.

In one aspect, the invention provides a hybrid polypeptide, comprising at least six different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 30 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least M5, M6, M14, M19, M24, and M29. In other embodiments, the polypeptide elicits an immune response against more than one antigen of group A streptococci comprising at least M2, M11, M22, M33, M43, M59, and M94, or at least M75, M76, M77, M89, M92, M101, and M114, or at least Spa, M1.0,

M1.2, M3, M12, M18, and M28, or against 27 or more antigens of group A streptococci. In more embodiments, the amino-terminal immunogenic peptide of the hybrid polypeptide is M24, M1, M2, or M89. In still other embodiments, the polypeptides are recombinant and the immunogenic peptides are linked in tandem, the polypeptides having a structure of M24-M5-M6-M19-M29-M14-M24, M2-M43-M94-M22-M11-M59-M33-M2, M89-M101-M77-M114-M75-M76-M92-M89, or M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0. In yet other embodiments, the hybrid polypeptides comprise the amino acid sequence of SEQ ID NOS:2, 4, 6, or 8. In other embodiments, the immunogenic peptides are linked by at least two amino acids encoded by a nucleic acid sequence that is a restriction enzyme recognition site, wherein the recognition site includes at least one of *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Nco*I, *Nhe*I, *Pml*I, *Pst*I, *Sal*I, and *Xho*I. In further embodiments, any of the aforementioned hybrid polypeptides capable of eliciting at least one opsonic antibody that is not a tissue cross-reactive antibody in a subject, wherein the subject is a human or an animal. In more embodiments, any of the aforementioned hybrid polypeptides further comprising a carboxy-terminal tag, wherein the carboxy-terminal tag is selected from the group consisting of alkaline phosphatase, β-galactosidase, hexahistidine, FLAG®, XPRESS®, and GST. In still more embodiments, any of the aforementioned hybrid polypeptides or fusion proteins further comprise at least one additional carboxy-terminal amino acid, wherein the additional carboxy-terminal amino acid is a D-amino acid or cysteine.

In another aspect, the invention provides any of the aforementioned hybrid polypeptides or fusion proteins wherein the polypeptides are synthetic. In certain embodiments, the synthetic hybrid polypeptides have one or more amino acids altered to a corresponding D-amino acid or are linked to an alkane such as acrylamide or an analog or derivative thereof. In a further embodiment, the synthetic hybrid polypeptides have immunogenic peptides linked to form a lysine core-branched peptide.

In still another aspect, the invention includes a nucleic acid molecule comprising a sequence encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8. In a related embodiment, there is provided a nucleic acid expression construct comprising an

expression control sequence operably linked to a polynucleotide encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8. In other embodiments, the expression construct comprises a nucleic acid expression vector selected from the group comprising a plasmid, phagemid, shuttle vector, cosmid, or virus. In one embodiment, the vector is plasmid pT5 (SEQ ID NO:17). In still another embodiment, there is provided a host cell containing any of the aforementioned nucleic acid constructs. In yet other embodiments, the host cell is selected from a bacterium, a yeast cell, a nematode cell, an insect cell, or a mammalian cell. In one embodiment, the host cell is the bacterium *Escherichia coli.*

In a further aspect, the present invention provides a plurality of antibodies, comprising two or more different antibodies wherein each antibody is specific for a different immunogenic peptide of a hybrid polypeptide, the polypeptide comprises at least six different immunogenic peptides linked in tandem, each peptide comprises at least 30 amino acids and the amino-terminal peptide is reiterated as a carboxy-terminal peptide, wherein the polypeptide is capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least M5, M6, M14, M19, M24, and M29. In other embodiments, the plurality of antibodies are specific for more than one antigen of group A streptococci comprising at least M2, M11, M22, M33, M43, M59, and M94, or at least M75, M76, M77, M89, M92, M101, and M114, or at least Spa, M1.0, M1.2, M3, M12, M18, and M28, or specific for 27 or more antigens of group A streptococci. In yet other embodiments, any of the aforementioned antibodies wherein the hybrid polypeptides comprise the amino acid sequence of SEQ ID NOS:2, 4, 6, or 8. In other embodiments, any of the aforementioned antibodies are opsonic and not tissue cross-reactive in a subject. In one embodiment, the antibodies are polyclonal.

In yet another aspect, there is provided a method of producing a hybrid polypeptide, comprising culturing a host cell containing any of the aforementioned nucleic acid expression vectors comprising at least one expression control sequence operably linked to a nucleic acid molecule encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8, under conditions and for a time sufficient for expression of the polypeptide. In a related

aspect, the invention provides any of the aforementioned hybrid polypeptides produced by the aforementioned method.

In another aspect, there is provided a composition comprising a pharmaceutically acceptable carrier and any of the aforementioned hybrid polypeptides. In other aspects, the invention is a cocktail composition comprising a pharmaceutically acceptable carrier and a mixture of at least two or three of any of the aforementioned hybrid polypeptides. In one embodiment, the cocktail compositions include at least one of the hybrid polypeptides having a Spa immunogenic peptide. In other embodiments, provided are any of the aforementioned compositions wherein the hybrid polypeptides are mixed in equimolar amounts. In further embodiments, any of the aforementioned compositions further comprise an adjuvant such as alum or Freund's.

In still another aspect, the invention provides a method for preventing a microbial infection, comprising administering to a subject any of the aforementioned compositions at a dose sufficient to elicit antibodies specific for one or more hybrid polypeptide, wherein the antibodies are opsonic and are not tissue cross-reactive. In certain embodiments, the microbial infection being prevented is a streptococcal infection and in a related embodiment is a group A streptococcal infection. In some embodiments, any of the aforementioned compositions are administered to a subject by a route selected from enteral, parenteral, transdermal, transmucosal, or inhalation. In further embodiments, the compositions are administered to a human or animal subject and the compositions further comprise an adjuvant such as alum or Freund's. In another related aspect, there is provided isolated antibodies produced by the aforementioned methods for preventing a microbial infection. In certain embodiments, the antibodies produced by these methods will comprise at least one antibody specific for a M serotype not represented in a hybrid polypeptide, such as M4. In still another embodiment, there is provided a method for treating or preventing a microbial infection comprising administering to a subject a composition comprising a pharmaceutically acceptable carrier and any of the aforementioned primary of antibodies. In one embodiment, the subject is an animal or human.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic diagram of the four hybrid polypeptides used as a vaccinating agent. The oligonucleotide primers used to amplify a 5' *emm* gene fragment (*i.e.*, immunogenic peptide) by PCR for each serotype were synthesized to include the indicated unique restriction enzyme sites, which encode amino acids that link the immunogenic peptides in tandem (indicated by the lines between the boxes). Each box represents an amino-terminal M protein immunogenic peptide, wherein the number within each box designates the serotype and the number below each box indicates the number of amino acids comprising each immunogenic peptide. Serotype M101 was formerly designated *stNS5*, serotype M114 was formerly designated *st2967*, and serotype M94 was formerly designated M13W and M13. The "6xHis" refers to the presence of a hexahistidine, which is optional.

Figures 2A to 2C show the nucleotide and amino acid sequence of hybrid polypeptide Hexavalent A.3. The restriction enzyme sites and the beginning point of each M serotype immunogenic peptide are indicated. Eleven Arg codons were optimized for expression in *E. coli* by mutating AGG/AGA codons to CGT/CGC codons (mutated bases are underlined).

Figures 3A to 3D show the nucleotide and amino acid sequence of hybrid polypeptide Hexavalent B.3a. The restriction enzyme sites and the beginning point of each M serotype immunogenic peptide are indicated. Nine Arg codons were optimized for expression in *E. coli* by mutating AGG/AGA codons to CGT/CGC codons (mutated bases are underlined).

Figures 4A to 4C show the nucleotide and amino acid sequence of hybrid polypeptide Hexavalent C.2. The restriction enzyme sites and the beginning point of each M serotype immunogenic peptide are indicated. Seven Arg codons were optimized for expression in *E. coli* by mutating AGG/AGA codons to CGT codons (mutated bases are underlined).

Figures 5A to 5C show the nucleotide and amino acid sequence of hybrid polypeptide Hexavalent D.3. The restriction enzyme sites and the beginning point of each

M serotype immunogenic peptide are indicated. Five Arg codons were optimized for expression in *E. coli* by mutating AGG/AGA codons to CGT/CGC codons (mutated bases are underlined).

Figure 6 shows a summary of the 28 most common invasive group A streptococci M serotypes isolated in the U.S. between August 12, 2000 and July 16, 2001. The data were part of ongoing studies conducted by the Active Bacterial Core Surveillance Program of the Centers for Disease Control and Prevention. These 28 serotypes accounted for 92.1% of the 3,424 invasive isolates submitted during this period.

Figure 7 shows that type-specific M protein and Spa antibodies were elicited by a vaccinating agent comprising a cocktail of four different hybrid polypeptides (*i.e.*, a 27-valent vaccine) when examined by ELISA. Each bar represents serum from one rabbit.

Figure 8 shows the results of *in vitro* opsonization assays using immune sera, which were elicited in rabbits with a composition comprising a cocktail of four different hybrid polypeptides (*i.e.*, a 27-valent vaccine). A positive response was considered at least a 3-fold increase over opsonization with pre-immune (*i.e.*, 30% or greater opsonization). The pre-immune sera resulted in less than 10% opsonization of each streptococcal serotype. Each bar represents serum from one rabbit.

Figure 9 shows the results of bactericidal activity assays using immune sera, which were elicited in rabbits with a vaccinating agent comprising a cocktail of four different hybrid polypeptides (*i.e.*, a 27-valent vaccine). A positive response was considered at least a 50% percent killing. Each bar represents serum from one rabbit.

Figure 10 shows that an immune response with bactericidal activity is elicited against group A streptococcal serotype 4, which is a serotype not represented in the 27-valent vaccinating agent. Each bar represents serum from one rabbit. Immune sera were from a rabbit immunized at 0, 4, and 8 weeks (stripped bar) and from a rabbit immunized at 0, 4, and 16 weeks (open bar). The serotype 4 streptococcal isolates were from five different geographic locations, including Florida (FL); Illinois (IL); California (CA); Connecticut (CT); and South Dakota (SD).

Figure 11 shows the geometric mean antibody titers before (Day 0) and after (Day 134) immunization of human subjects, as determined by ELISA. The graph is in a $\log_{10}$ scale. Serotype M101 was formerly designated *stNS5*, serotype *st*2967 is now designated M114, and serotype M13 is now designated M94.

Figure 12 shows the natural log fold-increase in ELISA-determined antibody titer (bars) as compared to percent increase in Group A streptococcal killing by polymorphonuclear cells in 26 serotypes. Serotype M101 was formerly designated *stNS5*, serotype M114 was formerly designated *st*2967, and serotype M13 is now designated M94.

Figure 13 shows the fold-increase in geometric mean antibody titers in human subjects immunized with a hexavalent polypeptide (Hexa 1.2) as compared to the geometric mean antibody titers in human subjects immunized with a 27-valent cocktail of four hybrid multivalent polypeptides (Hexa A.3, Septa B.3a, Septa C.2, and Septa D.3; *see* Figure 1). Geometric mean antibody titers were calculated based on ELISAs against the six serotypes present in both vaccines.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to hybrid polypeptides of streptococcal antigens and compositions thereof, which are capable of eliciting protective antibodies against streptococci. Furthermore, the compositions may include a single hybrid polypeptide or a combination of several different hybrid polypeptides, which may be useful to elicit an immune response against group A streptococci. In one aspect, one or more hybrid polypeptide may be formulated as a composition for simultaneously treating or preventing an infection by several different group A streptococcal serotypes, as well as treating or preventing infection by serotypes not represented in the hybrid polypeptide(s). The present invention also provides isolated nucleic acids that encode such hybrid polypeptides, as well as methods of expressing and purifying such hybrid polypeptides. A hybrid polypeptide, as described herein, comprises several different linked immunogenic peptides wherein each peptide comprises, for example, an amino-terminal portion of a streptococcal M protein and can elicit opsonic

antibodies specific for a particular group A streptococcal serotype without eliciting antibodies that are cross-reactive with host tissue.

The present invention also provides antibodies specific for each immunogenic peptide serotype included in a hybrid polypeptide or in a combination of hybrid polypeptides, as well as antibodies specific for serotypes not included in the hybrid polypeptide(s) (*i.e.*, antibodies that confer cross-protection between serotypes). The invention, therefore, relates generally to the surprising discovery that highly complex, multivalent hybrid polypeptide-based vaccines are feasible to simultaneously elicit broadly protective antibodies against several different streptococcal serotypes. Moreover, a composition comprising a cocktail of more than one multivalent hybrid polypeptide used as a vaccine unexpectedly elicits a much more robust antibody response in humans than does a composition comprising a single multivalent hybrid polypeptide, wherein the increase in antibody response also results in an increase in antibody function (*i.e.*, increased ability to opsonize and/or to kill microorganisms). As used herein, a "cocktail" of hybrid polypeptides refers to a composition comprising at least two different hybrid polypeptides of this invention. Hence, the invention also relates generally to the surprising discovery that a cocktail of multivalent hybrid polypeptides can function synergistically to elicit a greater antibody response that is a protective immune response. Accordingly, the compositions and methods of the present invention may be readily used to treat or prevent streptococcal infections. Discussed in more detail below are hybrid polypeptides and assorted compositions thereof (including admixtures or cocktails) suitable for use within the present invention, as well as exemplary methods of making such hybrid polypeptides and compositions, and therapeutic uses thereof.

HYBRID POLYPEPTIDES

The present invention is directed generally to multivalent immunogenic hybrid polypeptides of M proteins and M-like proteins, including fusions to other proteins. The immunogenic M and M-like peptides may comprise any portion of an M protein that is immunogenic, which may or may not confer serotype specificity. A plurality of different

multivalent immunogenic hybrid polypeptides can be mixed or combined into a cocktail composition for use in eliciting a protective immune response. The present invention further provides methods for producing synthetic or recombinant multivalent immunogenic hybrid M polypeptides, including fusion proteins. For example, host cells containing hybrid polypeptide-encoding nucleic acid expression constructs may be cultured to produce recombinant hybrid polypeptides. Also contemplated are methods for treating or preventing a microbial infection or eliciting an immune response using a hybrid polypeptide or a combination of hybrid polypeptides (including fusion proteins).

By way of background and not wishing to be bound by theory, streptococcal species are Gram-positive bacteria that are grouped based on immunological differences in their cell wall polysaccharides and are designated, for example, group A, B, C, F, and G. Specifically, group A streptococci (GrAS) are clinically important microorganisms that colonize the throat and skin. GrAS are responsible for a variety of suppurative infections (*e.g.*, strep throat, necrotizing fasciitis) and non-suppurative sequelae (*e.g.*, acute rheumatic fever, reactive arthritis) (*see generally* Cunningham, *Clin. Microbiol. Rev. 13*:470, 2000). GrAS have two major, surface-exposed anti-phagocytic factors, capsule and M protein, which allow these organisms to colonize and survive in a host. The M protein, which is encoded by the *emm* gene, extends from the cell surface as an α-helical coiled-coil dimer that appears as a fibril on the surface of GrAS. The M proteins are a diverse group, which have been serologically separated into M protein serotypes.

Currently, more than 120 M protein serotypes have been identified, and within some serotypes there have been identified several subtypes. For example, without limitation, type 1 M protein has related subtypes 1.1-1.15, and type 12 has subtypes 12.1-12.9. In addition, as is known in the art, unclassified serotypes may have an initial designation (such as *st*2967) and ultimately receive a "final" classification (such as *st*2967 now being classified as M type 114), or previously classified serotypes may be reclassified with a different number (such as M13W is now reclassified as M54). In certain embodiments, amino-terminal portions of any one or more of known M protein from serotypes 1-120+, or from unknown serotypes, can be used to generate immunogenic

peptides for inclusion in multivalent immunogenic hybrid polypeptides of the instant invention. Any numerical ranges recited herein are to be understood to include any integer within the range and, where applicable (*e.g.*, concentrations), fractions thereof, such as one tenth and one hundredth of an integer (unless otherwise indicated).

Furthermore, the M protein is part of a superfamily of proteins, which includes without limitation, immunoglobulin binding proteins (*e.g.*, FrcA), M-like proteins (*e.g.*, Mrp and Spa), and M proteins. Accordingly, as used herein, "M protein" refers to the superfamily of proteins, which includes any known or unknown M protein (with or without a designated serotype), as well as M-like proteins, such as Spa (*see, e.g.*, Dale *et al.*, *J. Clin. Invest. 103*:1261, 1999 and McLellan *et al.*, *Infect. Immun. 69*:2943, 2001), Mrp (*see, e.g.*, Boyle *et al.*, *J. Infect. Dis. 177*:991, 1998), immunoglobulin binding proteins (*see, e.g.*, Podbielski *et al.*, *Mol. Microbiol. 12*:725, 1994; Whatmore and Kehoe, *Mol. Microbiol. 11*:363, 1994), and the like. As described herein and is understood in the art, the serotype of an M protein may be reclassified and result in a change to a different type, a new type, or even a subtype. Also, as used herein, "M protein serotype" or "M type" refers to all M proteins within that serotype, including all subtypes, related proteins, and the like. Furthermore, reference to a particular M type may be applied interchangeably as follows, by way of illustration and not limitation, serotype 1 M protein, type 1, M1, and the like, which as set forth above, includes all subtypes as well.

As described above, group A streptococci have developed a system for avoiding some of the antimicrobial defenses of a human host. Strains of streptococci that express capsule and M protein can evade phagocytosis by, for example, polymorphonuclear leukocytes or neutrophils and multiply in a host that has not been exposed to streptococci (*i.e.*, have non-immune blood). Yet, a subject may develop resistance to a streptococcal infection if the host can produce protective antibodies directed against streptococci. Protection against GrAS generally correlates with the presence of opsonizing antibody against type-specific M protein (*see, e.g.*, Lancefield, *J. Immunol. 89*:307, 1962), and the development of secretory or mucosal antibodies is suspected of playing an important role in preventing initial colonization by streptococci. However, the

pathogenesis of some of the non-suppurative sequelae may be due to host tissue cross-reaction with streptococcal antibodies. As used herein, "tissue cross-reactivity" means a host antigen shares at least one epitope with a foreign antigen, such as a streptococcal antigen. For example, the different antigens may share an identical amino acid sequence, may be homologous but non-identical, or may be dissimilar molecules with a shared epitope (*e.g.*, protein and carbohydrate or protein and nucleic acid molecule). Thus, an effective vaccine against streptococcal infections would preferably elicit an immune response that includes antibodies that are not tissue cross-reactive (*i.e.*, do not cause autoimmune disease) and are protective (*i.e.*, opsonic) against many of the clinically important serotypes.

The present invention pertains to hybrid polypeptides or variants thereof having a plurality of immunogenic peptides from M or M-like proteins, or nucleic acid molecules encoding such polypeptides or variants thereof. As used herein, "immunogenic peptide" refers to any streptococcal peptide or polypeptide having at least one epitope capable of eliciting an immune response, which are the component units of the hybrid polypeptides. Preferably, the epitope is within an amino-terminal portion of a streptococcal M protein or Spa protein and more preferably is an opsonic epitope that does not elicit tissue cross-reactive antibodies.

The present invention also provides a rational vaccine design approach for selection of the streptococcal M protein serotypes to be included in the multivalent hybrid polypeptide vaccine. Some criteria that may be used, without limitation, include identifying the M protein serotypes that are frequent causes of uncomplicated pharyngitis, identifying the serotypes that are commonly recovered from normally sterile sites (*i.e.*, invasive strains) (*e.g.*, useful data may obtained from the Active Bacterial Core Surveillance of the Emerging Infections Program Network, supported by the Centers for Disease Control and Prevention; *see also* Beall *et al.*, *J. Clin. Microbiol.* *35*:1231, 1997; Schuchat *et al.*, *Emerg. Infect. Dis.* 7:92, 2001), and identifying serotypes that are considered currently or historically "rheumatogenic" (Bisno AL. The concept of rheumatogenic and nonrheumatogenic group A streptococci. In: Reed SE, and J. B.

Zabriskie, ed. Streptococcal Diseases and the Immune Response. New York: Academic Press, 1980:789-803). Also useful is the amino-terminal peptide fragment of Spa, a new protective antigen that is expressed by several serotypes of group A streptococci (Dale *et al., J. Clin. Invest. 103*:1261, 1999), or any other streptococcal antigen capable of eliciting protective antibodies may be used.

As described herein and known in the art, M protein amino acid sequences selected for inclusion in a hybrid polypeptide are available at the CDC *emm* typing center website (www.cdc.gov/ncidod/biotech/strep/emmtypes.html). In addition, to eliminate the possibility of eliciting tissue cross-reactive antibodies, the amino-terminal regions of a mature M protein and Spa may be compared to known human proteins to detect any homology (*e.g.*, using BlastP). Preferably, amino-terminal M protein portions having five or more contiguous amino acid matches with a human protein are excluded. In addition, the selected regions of the M peptides and Spa are preferably analyzed by the method of Hopp and Woods (*Mol. Immunol. 20*:483, 1983) to ensure the integrity of hydrophilic peaks. For example the Hopp and Woods method may be helpful in predicting that a particular immunogenic peptide from an M protein amino-terminal may best be fused with a certain subset of other immunogenic peptides and not with others.

In preferred embodiments, the hybrid polypeptides or variants, and combinations thereof, may be designed to be vaccines specific for streptococci associated with, for example, pharyngitis, scarlet fever, acute rheumatic fever, necrotizing faciitis, cellulitis, meningitis, pneumonia, toxic shock syndrome, bacteremia, septicemia, septic arthritis, pyoderma, skin infections, impetigo, erysipelas, soft-tissue infection, nephritis, pyrogenic reactions, and the like. Additionally, the vaccines may be designed to treat or prevent streptococcal infections in particular populations (*e.g.*, immunocompromised patients, children, and elderly) particular geographic populations (*e.g.*, Australian aborigines), and particular geographic locations (*e.g.*, temperate regions or Scandinavian countries). Preferably, the amino-terminal portions of the different M proteins that comprise an immunogenic peptide can elicit opsonic antibodies that do not cross-react with host tissue.

In another preferred aspect, the M protein serotype is selected based on its prevalence on the most common streptococci known to currently be associated with a particular disease (*e.g.*, serotypes associated with pharyngitis or skin infections) or sequelae. In a further preferred aspect, the instant invention may be used to design and generate a variety of different multivalent immunogenic hybrid polypeptides that may be directed to, or be admixed in particular groupings to address, shifts in prevalent streptococci. For example, as is known in the art, the most prevalent streptococci serotype associated with a disease today, such as ARF, may not be as prevalent or important 5, 10, or 15 years. In another example, a particular streptococcal serotype may be prevalent in Europe and suddenly become important in South America. Hence the immunogenic peptides that comprise a multivalent hybrid polypeptide may be changed, or different multivalent hybrid polypeptides may be mixed and matched to create a desired cocktail, designed for use in a particular population or location to attack the most clinically relevant streptococci as needed.

In certain preferred embodiments, an immunogenic peptide comprises an amino-terminal portion of a M protein or Spa protein having 10-70 amino acids, or any integer in that range; more preferably having 20-65 amino acids, or any integer in that range; and most preferably 30-60 amino acids, or any integer in that range. In particularly preferred embodiments, a hybrid polypeptide comprises at least six or seven different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 30 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one serotype of group A streptococci. Preferably, a hybrid polypeptide is capable of eliciting an immune response against at least serotypes 5, 6, 14, 19, 24, and 29; or at least against serotypes 2, 11, 22, 33, 43, 59, and 94; or at least against serotypes 75, 76, 77, 89, 92, 101, and 114; or at least against serotypes 1.0, 1.2, 3, 12, 18, and 28. In another preferred embodiment, a hybrid polypeptide is capable of eliciting an immune response against a M serotype not represented in the hybrid polypeptide, such as serotype 4.

In certain embodiments, the hybrid polypeptides have at least 50% to 100% amino acid identity, or any integer in that range, to the amino acid sequence as set forth in SEQ ID NOS:2, 4, 6, and 8; preferably 60%-99% identity or any integer in that range, more preferably 70%-97% identity or any integer in that range, and most preferably 80%-95% identity or any integer in that range. As used herein, "percent identity" or "% identity" is the percentage value returned by comparing the whole of the subject polypeptide, peptide, or variant thereof sequence to a test sequence using a computer implemented algorithm, typically with default parameters. Sequence comparisons can be performed using any standard software program, such as BLAST, tBLAST, pBLAST, or MegAlign. Still others include those provided in the Lasergene bioinformatics computing suite, which is produced by DNASTAR® (Madison, Wisconsin). References for algorithms such as ALIGN or BLAST may be found in, for example, Altschul, *J. Mol. Biol. 219*:555-565, 1991; or Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA 89*:10915-10919, 1992. BLAST is available at the NCBI website (www.ncbi.nlm.nih.gov/BLAST). Other methods for comparing multiple nucleotide or amino acid sequences by determining optimal alignment are well known to those of skill in the art (*see, e.g.*, Peruski and Peruski, *The Internet and the New Biology: Tools for Genomic and Molecular Research* (ASM Press, Inc. 1997); Wu *et al.* (eds.), "Information Superhighway and Computer Databases of Nucleic Acids and Proteins," in Methods in Gene Biotechnology, pages 123-151 (CRC Press, Inc. 1997); and Bishop (ed.), Guide to Human Genome Computing, 2nd Edition, Academic Press, Inc., 1998). As used herein, "similarity" between two peptides or polypeptides is generally determined by comparing the amino acid sequence of one peptide or polypeptide to the amino acid sequence and conserved amino acid substitutes thereto of a second peptide or polypeptide. Fragments or portions of the hybrid polypeptides of the present invention may be employed for producing the corresponding full-length hybrid polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length hybrid polypeptides. Similarly, fragments or portions of the nucleic acids of the present invention may be used to synthesize full-length nucleic acids of the present invention.

The hybrid polypeptides and corresponding nucleic acids of the present invention are preferably provided in an isolated form, and in certain preferred embodiments, are purified to homogeneity. As used herein, the term "isolated" means that the material is removed from its original or natural environment. For example, a naturally occurring nucleic acid molecule or polypeptide present in a living animal or cell is not isolated, but the same nucleic acid molecule or polypeptide is isolated when separated from some or all of the co-existing materials in the natural system. The nucleic acid molecules, for example, could be part of a vector and/or such nucleic acids or polypeptides could be part of a composition and still be isolated in that such vector or composition is not part of its natural environment.

The present invention also pertains to hybrid polypeptides and variants thereof produced synthetically or recombinantly, and preferably recombinantly. The immunogenic peptide components of the hybrid polypeptides may be synthesized by standard chemical methods, including synthesis by automated procedure. In general, immunogenic peptides are synthesized based on the standard solid-phase Fmoc protection strategy with HATU as the coupling agent. The immunogenic peptide is cleaved from the solid-phase resin with trifluoroacetic acid containing appropriate scavengers, which also deprotects side chain functional groups. Crude immunogenic peptide is further purified using preparative reversed-phase chromatography. Other purification methods, such as partition chromatography, gel filtration, gel electrophoresis, or ion-exchange chromatography may be used. Other synthesis techniques known in the art may be employed to produce similar immunogenic peptides, such as the tBoc protection strategy, use of different coupling reagents, and the like. In addition, any naturally occurring amino acid or derivative thereof may be used, including D- or L-amino acids and combinations thereof. In particularly preferred embodiments, a synthetic hybrid polypeptide of the invention will have a M2, M12, M24, or M89 immunogenic peptide at the amino-terminus.

As described herein, the hybrid polypeptides of the invention may be recombinant, wherein the hybrid polypeptide is expressed from a polynucleotide encoding a desired hybrid polypeptide that is operably linked to an expression control sequence (*e.g.*,

promoter) in a nucleic acid expression construct. In particularly preferred embodiments, a recombinant hybrid polypeptide of the invention will have a M2, M12, M24, or M89 immunogenic peptide at the amino-terminus. Some preferable recombinant hybrid polypeptides comprise immunogenic peptides linked in tandem having the structure of M24-M5-M6-M19-M29-M14-M24, M2-M43-M94-M22-M11-M59-M33-M2, M89-M101-M77-M114-M75-M76-M92-M89, or M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0, and any combination thereof. Most preferably, a hybrid polypeptide comprises the amino acid sequence of SEQ ID NOS:2, 4, 6, or 8, and variants thereof. As set forth above and herein, any M serotype may be included in the present invention, preferably serotypes 1.0, 1.2, 2, 3, 5, 6, 11, 12, 14, 18, 19, 22, 24, 28, 29, 33, 43, 59, 75, 76, 77, 89, 92, 94, 101, and 114 are included in the hybrid polypeptides. In certain preferred embodiments as described herein, hybrid polypeptides of the subject invention capable of eliciting at least one opsonic antibody that is not a tissue cross-reactive antibody in a subject, wherein the subject is an animal or a human.

In another preferred embodiment, the hybrid polypeptides are linked by at least two amino acids encoded by a nucleic acid sequence that is a restriction enzyme recognition site, wherein the restriction sites may be any one or more of *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Nco*I, *Nhe*I, *Pml*I, *Pst*I, *Sal*I, *Xho*I, and the like. Additional amino acid linkers may also be added synthetically as described herein. Preferably, the additional amino acids do not create any identity in sequence within a five amino acid stretch of a human protein. In addition, the hybrid polypeptides of the subject invention may further comprise at least one additional carboxy-terminal amino acid, wherein the additional amino acid is a D- or an L-amino acid. Any of the twenty naturally occurring amino acids or derivatives thereof may be added, such as cysteine, histidine, leucine, and glutamic acid. For example, the addition of cysteine may be useful to attach other constituents, such as a lipid, a carrier protein, and the like.

As described herein, the invention also provides hybrid polypeptide fusion proteins comprising hybrid polypeptides fused to an additional functional or non-functional polypeptide sequence that permits, for example by way of illustration and not limitation,

detection, isolation and/or purification of the hybrid polypeptide fusion proteins. For instance, an additional functional polypeptide sequence may be a tag sequence, which includes fusion proteins that may in certain embodiments be detected, isolated and/or purified by protein-protein affinity (*e.g.*, receptor-ligand), metal affinity or charge affinity methods. In certain other embodiments the hybrid polypeptide fusion proteins may be detected by specific protease cleavage of a fusion protein having a sequence that comprises a protease recognition sequence, such that the hybrid polypeptides may be separable from the additional polypeptide sequence. In addition, the hybrid polypeptides may be made synthetically including additional amino acids, a carrier protein, or a tag sequence, which may be located at either the amino- or carboxy-terminal end. In particularly preferred embodiments, for example, recombinant hybrid polypeptides are fused in-frame to a carboxy-terminal tag, which tag may be any one of alkaline phosphatase, β-galactosidase, hexahistidine (6xHis), FLAG® epitope tag (DYKDDDDK, SEQ ID NO:18), or GST, and the like. Most preferred are hybrid polypeptide fusion proteins that facilitate affinity detection and isolation of the hybrid polypeptides and may include, for example, poly-His or the defined antigenic peptide epitopes described in U.S. Patent No. 5,011,912 and in Hopp et al., (1988 *Bio/Technology 6*:1204), or the XPRESS™ epitope tag (DLYDDDDK, SEQ ID NO:19; Invitrogen, Carlsbad, CA). The affinity sequence may be a hexa-histidine tag as supplied by a vector. For example, a pBAD/His (Invitrogen) or a pQE-30 (Qiagen, Valencia, CA) vector can provide a polyhistidine tag for purification of the mature protein fusion from a particular host, such as a bacterium. Alternatively, the affinity sequence may be added either synthetically or engineered into the primers used to recombinantly generate the nucleic acid sequence (*e.g.*, using the polymerase chain reaction) encoding an immunogenic peptide of the multivalent hybrid polypeptide. Preferably, a multivalent hybrid polypeptide is fused to a polyhistidine and is encoded by a recombinant nucleic acid sequence encoding such a fusion protein.

The immunogenic peptides may also be chemically linked to form the desired hybrid polypeptides, including additional amino acid sequences, by a variety of methods, as provided herein and known in the art (*see, generally,* Jackson *et al.*, *Vaccine*

18:355, 2000). Recombinant or synthetic peptides may be linked to form linear (*see, e.g.*, Leclerc, *et al.*, *Eur. J. Immunol.* 17:26, 1987 and Francis, *et al.*, *Nature* 330:168, 1987) or branched (*see, e.g.*, Fitzmaurice, *et al.*, *Vaccine* 14:553, 1996) constructs, or may be linked using chemical ligation of epitopes (*see, e.g.*, Tam and Spetzler, *Biomed. Pept. Proteins Nucleic Acids* 1:123, 1995; Rose, *J. Am. Chem. Soc.* 116:30, 1994; and Dawson, *et al.*, *Science* 266:776, 1994). Peptides may also be linked via the multiple antigen peptide system to form branched hybrid polypeptides (*see, e.g.*, Tam, *Proc. Natl. Acad. Sci. USA 85*:5409, 1988; U.S. Patent No. 5,229,490). The multiple antigen peptide system makes use of multifunctional core molecules where each functional group on the core molecule forms at least two branches, the principal units of which are also multifunctional. For example, lysine may be used as the core peptide because it has one carboxyl functional group and two (α and ε) amine functional groups. Each multifunctional unit in a branch provides a base for added growth, resulting in exponential growth of the dendritic polymer. Peptides may then be joined to the dendritic core using a linking molecule (*e.g.*, glycine). The multiple antigen peptide system links a large number of synthetic peptides to the functional group of a dendritic core molecule providing a high concentration of synthetic peptides in a low molecular volume. Preferably, either two or three levels of geometrically branched lysines are used, wherein these lysine cores form a tetrameric and octameric core structure, respectively. The multiple antigen peptide system may also include a lipophilic anchoring moiety attached to the core molecule, thereby eliminating the need for an adjuvant formulated in a peptide vaccine otherwise requiring one for immunostimulation (*see, e.g.*, U.S. Patent No. 5,580,563). In one preferred embodiment, a hybrid polypeptide of the invention comprises immunogenic peptides linked to form a lysine core-branched polypeptide.

Additionally, similar or different synthetic peptides may be linked by controlled polymerization through derivatization of the amino-terminus of a peptide with the acryloyl (CH_2=CH–) group using acryloyl chloride (*see, e.g.*, O'Brien-Simpson, *et al.*, *J. Am. Chem. Soc.* 119:1183, 1997; Jackson, *et al.*, *Vaccine* 15:1697, 1997). The derivatized peptides are then polymerized singly, or in admixture with similarly derivatized

peptides, by free radical initiation of chain elongation. As a result, peptides are assembled into polymers in which the peptide determinants form side chains pendant from an alkane backbone. The hybrid polypeptides and fusion proteins as described herein may be constructed as set forth above. In one preferred embodiment, a hybrid polypeptide of the invention comprises immunogenic peptides linked by an alkane backbone. In certain embodiments, the alkane backbone is acrylamide or an analog or derivative thereof.

THERAPEUTIC FORMULATIONS AND METHODS OF USE

The invention also relates to pharmaceutical compositions that contain one or more hybrid multivalent polypeptides, which may be used to elicit an immune response. The invention further relates to methods for treating and preventing microbial infections by administering to a subject a hybrid polypeptide or a mixture of hybrid polypeptides at a dose sufficient to elicit antibodies specific for one or more hybrid polypeptide, as described herein. A hybrid polypeptide or a cocktail of hybrid polypeptides is preferably part of a pharmaceutical composition when used in the methods of the present invention.

In certain aspects, the invention provides a composition comprising a pharmaceutically acceptable carrier, diluent, or excipient, and any of the multivalent hybrid polypeptides of the subject invention and any combination thereof. A preferred embodiment is a pharmaceutically acceptable carrier and a mixture of at least two or three hybrid polypeptides of the subject invention. In yet another preferred embodiment, a hybrid polypeptide that comprises at least seven different immunogenic peptides linked in tandem, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 50 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one serotype of group A streptococci comprising at least serotypes 1.0, 1.2, 3, 12, 18, and 28, is combined with at least one other hybrid polypeptide of the subject invention and a pharmaceutically acceptable carrier. In a more preferred embodiment, a composition comprises a pharmaceutically acceptable carrier and a mixture of the hybrid polypeptides of SEQ ID NOS:2, 4, 6, and 8 or variants

thereof, and in other embodiment these four polypeptides or variants thereof are mixed in equimolar amounts and at least one of the hybrid polypeptides has a Spa peptide. In other embodiments, SEQ ID NOS:2, 4, 6, and 8 or variants thereof are provided with a polyhistidine tag or further comprise at least one additional amino acid, such as cysteine. Each of these formulations may further comprise an adjuvant, such as alum or Freund's, and a diluent such as water or PBS. Further, therapeutic compositions of the present invention should preferably be stable for several months and capable of being produced and maintained under sterile conditions.

The pharmaceutical composition will include at least one of a pharmaceutically acceptable vehicle, carrier, diluent, or excipient, in addition to one or more hybrid multivalent polypeptide or fusion protein thereof and, optionally, other components. For example, pharmaceutically acceptable carriers suitable for use with a composition of a hybrid polypeptide or fusion protein thereof, or cocktail of two or more hybrid polypeptide or fusion protein thereof, may include, for example, a thickening agent, a buffering agent, a solvent, a humectant, a preservative, a chelating agent, an adjuvant, and the like, and combinations thereof. Exemplary adjuvants are alum (aluminum hydroxide, REHYDRAGEL®), aluminum phosphate, proteosome adjuvant (*see, e.g.*, U.S. Patent Nos. 5,726,292 and 5,985,284), virosomes, liposomes with and without Lipid A, Detox (Ribi/Corixa), MF59, or other oil and water emulsions type adjuvants, such as nanoemulsions (*see, e.g.*, U.S. Patent No. 5,716,637) and submicron emulsions (*see, e.g.*, U.S. Patent No. 5,961,970), and Freund's complete and incomplete. Pharmaceutically acceptable carriers for therapeutic use are well known in the pharmaceutical art, and as described herein and, for example, in *Remington's Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro, ed., 18th Edition, 1990) and in *CRC Handbook of Food, Drug, and Cosmetic Excipients*, CRC Press LLC (S.C. Smolinski, ed., 1992).

"Pharmaceutically acceptable salt" refers to salts of the compounds of the present invention derived from the combination of such compounds and an organic or inorganic acid (acid addition salts) or an organic or inorganic base (base addition salts).

The compounds of the present invention may be used in either the free base or salt forms, with both forms being considered as being within the scope of the present invention.

In addition, the pharmaceutical composition may further include a diluent such as water or phosphate buffered saline (PBS). Preferably, diluent is PBS with a final phosphate concentration ranges from about 0.1 mM to about 1 M, more preferably from about 0.5 mM to about 500 mM, even more preferably from about 1 mM to about 50 mM, and most preferably from about 2.5 mM to about 10 mM; and the final salt concentration ranges from about 100 mM to about 200 mM and most preferably from about 125 mM to about 175 mM. Preferably, the final PBS concentration is about 5 mM phosphate and about 150 mM salt (such as NaCl). In certain embodiments, any of the aforementioned pharmaceutical compositions comprising a cocktail of multivalent hybrid polypeptides of the instant invention are preferably sterile.

The compositions can be sterile either by preparing them under an aseptic environment and/or they can be terminally sterilized using methods available in the art. Many pharmaceuticals are manufactured to be sterile and this criterion is defined by the USP XXII <1211>. Sterilization in this embodiment may be accomplished by a number of means accepted in the industry and listed in the USP XXII <1211>, including gas sterilization, ionizing radiation or filtration. Sterilization may be maintained by what is termed asceptic processing, defined also in USP XXII <1211>. Acceptable gases used for gas sterilization include ethylene oxide. Acceptable radiation types used for ionizing radiation methods include gamma, for instance from a cobalt 60 source and electron beam. A typical dose of gamma radiation is 2.5 MRad. When appropriate, filtration may be accomplished using a filter with suitable pore size, for example 0.22 μm and of a suitable material, for instance Teflon®. The term "USP" refers to U.S. Pharmacopeia (*see* www.usp.org; Rockville, MD).

The present invention also pertains to methods for preventing a microbial infection, comprising administering to a subject a composition of the subject invention at a dose sufficient to elicit antibodies specific for one or more hybrid polypeptide, wherein the antibodies are preferably opsonic and are not tissue cross-reactive. In certain embodiments

an infection is a streptococcal infection, such as a group A streptococcal infection. A subject suitable for treatment with a hybrid polypeptide formulation may be identified by well-established indicators of risk for developing a disease or well-established hallmarks of an existing disease. For example, indicators of an infection include fever, pus, microorganism positive cultures, inflammation, and the like. Infections that may be treated with a hybrid polypeptide of the subject invention include, without limitation, those caused by or due to microorganisms, whether the infection is primary, secondary, opportunistic, or the like. Examples of microorganisms include Gram-positive bacteria, such as streptococci.

The pharmaceutical compositions that contain one or more hybrid polypeptides may be in any form that allows for the composition to be administered to a subject, such as a human or animal. For example, multivalent hybrid polypeptide compositions of the present invention may be prepared and administered as a liquid solution or prepared as a solid form (*e.g.*, lyophilized), which may be administered in solid form, or resuspended in a solution in conjunction with administration. The hybrid polypeptide composition is formulated so as to allow the active ingredients contained therein to be bioavailable upon administration of the composition to a subject or patient or bioavailable via slow release. Compositions that will be administered to a subject or patient take the form of one or more dosage units, where for example, a tablet may be a single dosage unit, and a container of one or more compounds of the invention in aerosol form may hold a plurality of dosage units. In certain preferred embodiments, any of the aforementioned pharmaceutical compositions comprising a hybrid polypeptide or cocktail of hybrid polypeptides of the invention are in a container, preferably in a sterile container.

In one embodiment, the therapeutic composition is administered orally, and a hybrid polypeptide or cocktail composition of the invention is taken up by cells, such as cells located in the lumen of the gut. Other typical routes of administration include, without limitation, enteral, parenteral, transdermal/transmucosal, and inhalation. The term enteral, as used herein, is a route of administration in which the agent is absorbed through the gastrointestinal tract or oral mucosa, including oral, rectal, and sublingual. The term

parenteral, as used herein, describes administration routes that bypass the gastrointestinal tract, including intraarterial, intradermal, intramuscular, intranasal, intraocular, intraperitoneal, intravenous, subcutaneous, submucosal, and intravaginal injection or infusion techniques. The term transdermal/transmucosal, as used herein, is a route of administration in which the agent is administered through or by way of the skin, including topical. The term inhalation encompasses techniques of administration in which an agent is introduced into the pulmonary tree, including intrapulmonary or transpulmonary. Preferably, the compositions of the present invention are administered intramuscularly.

Depending upon the application, the dose of hybrid polypeptide in the compositions will vary, generally, from about 10 μg to about 10 mg, preferably from about 100 μg to 1 mg, more preferably from about 150 μg to 500 μg, and most preferably from about 200 μg to about 400 μg, in combination with the biologically acceptable excipient, adjuvant, binder, and/or diluent, including any integer with the dosing range. As used herein, the term "about" or "consists essentially of" refers to ± 10% of any indicated structure, value, or range. Booster immunizations may be given at multiple times, at desired intervals ranging from about 2 weeks to about 24 weeks, preferably 2, 4 and 8 week intervals, and more preferably 2, 4, and 16 week intervals, and even more preferably 0, 4, and 24 week intervals to maximize the immune response.

The invention further provides a plurality of antibodies produced by the method for preventing a microbial infection that comprises administering to a subject a composition of the subject invention at a dose sufficient to elicit antibodies specific for one or more hybrid polypeptide, wherein the antibodies are opsonic and are not tissue cross-reactive. In one embodiment, the antibodies comprise at least one antibody specific for a M serotype not represented in a hybrid polypeptide, such as type 4 M protein. In another embodiment, a method for treating or preventing a microbial infection comprises administering to a subject a composition comprising a pharmaceutically acceptable carrier and a plurality of antibodies of the subject invention.

ANTIBODIES AND ASSAYS

In another aspect, the hybrid polypeptides and variants thereof of the present invention are utilized to elicit antibodies specific for at least one epitope present on the hybrid polypeptides provided herein. Accordingly, the present invention also provides such antibodies. In preferred embodiments the antibodies bind to specific protective epitopes present on a M protein. Within the context of the present invention, the term "antibodies" includes polyclonal antibodies, monospecific antibodies, monoclonal antibodies, anti-idiotypic antibodies, fragments thereof such as $F(ab')_2$ and Fab fragments, and recombinantly or synthetically produced antibodies. Such antibodies incorporate the variable regions that permit a monoclonal antibody to specifically bind, which means an antibody is able to selectively bind to a peptide produced from an *emm* or *spa* sequence of this invention. "Specific for" refers to the ability of a protein (*e.g.*, an antibody) to selectively bind a polypeptide or peptide encoded by an *emm* or *spa* nucleic acid molecule or a synthesized hybrid polypeptide of this invention. Association or "binding" of an antibody to a specific antigen generally involve electrostatic interactions, hydrogen bonding, Van der Waals interactions, and hydrophobic interactions. Any one of these or any combination thereof can play a role in the binding between an antibody and its antigen. Such an antibody generally associates with an antigen, such as M protein, with an affinity constant (K_a) of at least 10^4, preferably at least 10^5, more preferably at least 10^6, still more preferably at least 10^7 and most preferably at least 10^8. Affinity constants may be determined by one of ordinary skill in the art using well-known techniques (*see* Scatchard, *Ann. N.Y. Acad. Sci. 51*:660-672, 1949). The affinity of a monoclonal antibody or antibody can be readily determined by one of ordinary skill in the art (*see* Scatchard, *Ann. N.Y. Acad. Sci. 51*:660-672, 1949).

In addition, the term "antibody," as used herein, includes naturally occurring antibodies as well as non-naturally occurring antibodies, including, for example, single chain antibodies, chimeric, bifunctional and humanized antibodies, as well as antigen-binding fragments thereof. Such non-naturally occurring antibodies may be constructed using solid phase peptide synthesis, may be produced recombinantly, or may be obtained,

for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains (Huse *et al.*, *Science 246*:1275-1281 (1989)). These and other methods of making, for example, chimeric, humanized, CDR-grafted, single chain, and bifunctional antibodies are well known in the art (Winter and Harris, *Immunol. Today 14*:243, 1993; Ward *et al.*, *Nature 341*:544, 1989; Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, 1992; Borrabeck, *Antibody Engineering*, 2d ed., Oxford Univ. Press, 1995; Hilyard *et al.*, *Protein Engineering: A practical approach*, IRL Press, 1992).

In a preferred embodiment, a plurality of antibodies comprises two or more different antibodies wherein each antibody is specific for a different immunogenic peptide of a hybrid polypeptide, the polypeptide comprises at least six different immunogenic peptides linked in tandem, each peptide comprises at least 30 amino acids and the amino-terminal peptide is reiterated as a carboxy-terminal peptide, wherein the polypeptide is capable of eliciting an immune response against more than one serotype of group A streptococci. Preferably, a hybrid polypeptide is capable of eliciting an immune response against at least serotypes 5, 6, 14, 19, 24, and 29; or at least against serotypes 2, 11, 22, 33, 43, 59, and 94; or at least against serotypes 75, 76, 77, 89, 92, 101, and 114; or at least against serotypes 1.0, 1.2, 3, 12, 18, and 28. In another preferred embodiment, a hybrid polypeptide is capable of eliciting an immune response against a M serotype not represented in the hybrid polypeptide, such as serotype 4. More preferably the antibodies are elicited by the hybrid polypeptides of SEQ ID NOS:2, 4, 6, and 8, individually or in combination, and most preferably at least one antibody is opsonic and not tissue cross-reactive in a subject. As used herein, "opsonic" means any epitope that enhances phagocytosis of a cell or particle having the epitope. As commonly understood by those having ordinary skill in the art, "opsonic antibodies" are antibodies that facilitate phagocytic activity of a particle having the antigen, such as a bacterial cell. In yet another preferred embodiment, the antibodies elicited by the hybrid polypeptides of the invention are polyclonal.

Polyclonal antibodies can be readily generated by one of ordinary skill in the art from a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, turkeys, rabbits, mice, or rats. Briefly, the desired hybrid polypeptide or mixtures of hybrid polypeptides, or variants thereof are administered to immunize an animal through parenteral, intraperitoneal, intramuscular, intraocular, or subcutaneous injections. The immunogenicity of the hybrid polypeptide of interest may be increased through the use of an adjuvant, such as alum and Freund's complete or incomplete adjuvant. Following several booster immunizations over a period of weeks, small samples of serum are collected and tested for reactivity to the desired M peptide or Spa peptide. Particularly preferred polyclonal immune sera give a signal that is at least three times greater than background. Once the titer of the animal has reached a plateau in terms of its reactivity to the hybrid polypeptide, larger quantities of polyclonal immune sera may be readily obtained either by weekly bleedings or by exsanguinating the animal.

For example, the hybrid polypeptides of SEQ ID NOS:2, 4, 6, and 8 (*see* Figure 1) were purified, mixed in equimolar concentrations, and formulated with alum to contain 400 μg total protein and 750 μg alum in each 0.5 ml dose. Three rabbits each received three intramuscular injections of the hybrid polypeptide cocktail at either 0, 4, and 8 weeks or 0, 4, and 16 weeks, and immune serum was recovered at 18 weeks. The harvested sera were analyzed by ELISA, as described herein, using the purified recombinant dimeric peptide components of each of the larger vaccine polypeptides. Using ELISA (with purified recombinant dimeric immunogenic peptides of the hybrid molecule), the immune sera from rabbits immunized at 0, 4, and 16 weeks were found to contain high titers of antibodies against the vast majority of the M peptides and the Spa peptide contained in the hybrid polypeptides of the cocktail (Figure 3). In certain preferred embodiments, the polyclonal antibodies include those that are specific for group A streptococci of serotypes 5, 6, 14, 19, 24, and 29, or of serotypes 2, 11, 22, 33, 43, 59, and 94, or of serotypes 75, 76, 77, 89, 92, 101, and 114, or of serotypes 1.0, 1.2, 3, 12, 18, and 28.

Monoclonal antibodies may also be readily generated using well-known techniques (*see* U.S. Patent Nos. RE 32,011, 4,902,614, 4,543,439, and 4,411,993; *see also Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, Kennett, McKearn, and Bechtol (eds.), 1980, and *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Briefly, in one embodiment, a subject animal such as a rat or mouse is injected with a desired protein or peptide. If desired, various techniques may be utilized in order to increase the resultant immune response generated by the protein, in order to develop greater antibody reactivity. For example, the desired protein or peptide may be coupled to another carrier protein (such as ovalbumin, keyhole limpet hemocyanin (KLH), or *E. coli* labile toxin B subunit) or through the use of adjuvants (such as alum or Freund's complete and incomplete adjuvant) and the like.

Once suitable antibodies have been obtained, they may be isolated or purified by many techniques well known to those of ordinary skill in the art (*see Antibodies: A Laboratory Manual*, Harlow and Lane, *supra*). Suitable techniques include peptide or protein affinity columns, HPLC or RP-HPLC, purification on protein A or protein G columns, or any combination of these techniques. Within the context of the present invention, the term "isolated" as used to define antibodies or antibodies means "substantially free of other blood components."

Several assays are available as described herein to examine the activity of the antibodies elicited by the hybrid polypeptides of the subject invention. An exemplary assay is an opsonophagocytosis assay, which detects phagocytosis facilitated by the presence of opsonic antibodies present in test antisera. Briefly, the assay measures the amount of phagocytosis of selected bacterial cells by neutrophils after preincubating the cells in the presence or absence of antisera raised against, for example, hybrid polypeptide immunogens. Preincubation with the immune sera coats the cells with M protein reactive antibodies, some of which will be opsonic antibodies elicited from opsonic epitopes present on the M protein antigens. Preincubated, coated cells are then mixed with whole blood from an animal, typically a mammal for which opsonic protection is to be sought (*e.g.*, a

human) to determine the percentage of neutrophils that associate with the bacterial cells, which is a measure of phagocytic activity facilitated by opsonic antibodies. Immune sera containing opsonic antibodies induce a higher percentage of neutrophils associated with the selected bacteria than does immune sera lacking opsonic antibodies. In a variation of this test, the bactericidal activity of immune sera may be tested by incubating the immune sera with fewer bacterial cells, incubating in blood for a longer period of time, and then plating the mixture on a culture medium to score for viable bacteria. The presence of opsonic antibodies in the immune sera increase the number of bacteria destroyed by phagocytosis and, therefore, lowers the number of colony forming units (CFUs) detected on the plate culture. Another exemplary assay analyzes bactericidal activity of the test antibodies (*see* Example 5).

NUCLEIC ACID MOLECULES AND HOST CELLS

The invention also encompasses isolated nucleic acid molecules comprising a sequence encoding a hybrid polypeptide wherein each peptide comprises an amino-terminal portion of a streptococcal M protein (*e.g.*, SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16). Also provided by the present invention are nucleic acid expression constructs, and host cells containing such nucleic acids, which encode hybrid polypeptide and variants thereof, which hybrid polypeptides are capable of eliciting an immune response against more than one serotype of group A streptococci. This aspect of the invention pertains to isolated nucleic acid sequences encoding a hybrid polypeptide sequence as described herein, as well as those sequences readily derived from isolated nucleic acid molecules such as, for example, complementary sequences, reverse sequences and complements of reverse sequences.

"Nucleic acid" or "nucleic acid molecule" refers to any of deoxyribonucleic acid (DNA), ribonucleic acid (RNA), oligonucleotides, fragments generated by the polymerase chain reaction (PCR), and fragments generated by any of ligation, scission, endonuclease action, and exonuclease action. Nucleic acids may be composed of monomers that are naturally occurring nucleotides (such as deoxyribonucleotides and

ribonucleotides), analogs of naturally occurring nucleotides (*e.g.*, α-enantiomeric forms of naturally-occurring nucleotides), or a combination of both. Modified nucleotides can have modifications in sugar moieties and/or in pyrimidine or purine base moieties. Sugar modifications include, for example, replacement of one or more hydroxyl groups with halogens, alkyl groups, amines, and azido groups, or sugars can be functionalized as ethers or esters. Moreover, the entire sugar moiety may be replaced with sterically and electronically similar structures, such as aza-sugars and carbocyclic sugar analogs. Examples of modifications in a base moiety include alkylated purines and pyrimidines, acylated purines or pyrimidines, or other well-known heterocyclic substitutes. Nucleic acid monomers can be linked by phosphodiester bonds or analogs of such linkages. Analogs of phosphodiester linkages include phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoranilidate, phosphoramidate, and the like. The term "nucleic acid" also includes so-called "peptide nucleic acids," which comprise naturally occurring or modified nucleic acid bases attached to a polyamide backbone. Nucleic acids can be either single stranded or double stranded.

Further, an "isolated nucleic acid molecule" refers to a polynucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid construct, which has been separated from its source cell (including the chromosome it normally resides in) at least once in a substantially pure form. For example, a DNA molecule that encodes a Spa polypeptide, peptide, or variant thereof, which has been separated from a Streptococcus cell or from the genomic DNA of a Streptococcus cell, is an isolated DNA molecule. Another example of an isolated nucleic acid molecule is a chemically synthesized nucleic acid molecule. Nucleic acid molecules may be comprised of a wide variety of nucleotides, including DNA, cDNA, RNA, nucleotide analogues, or some combination thereof. In one preferred embodiment, an isolated nucleic acid molecule comprises a sequence encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16.

In certain aspects, the invention relates to nucleic acid vectors and constructs that include nucleic acid sequences of the present invention, and in particular to

"nucleic acid expression constructs" that include any polynucleotide encoding a hybrid polypeptide as provided above; to host cells that are genetically engineered with vectors and/or constructs of the invention and to the production and use in methods for treating or preventing a microbial infection or eliciting an immune response. The hybrid polypeptides may be expressed in mammalian cells, yeast, bacteria or other cells under the control of appropriate expression control sequences. Cell-free translation systems may also be employed to produce such proteins using RNAs derived from the nucleic acid expression constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described, for example, by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, NY, (1989), and may include plasmids, cosmids, shuttle vectors, viral vectors and vectors comprising a chromosomal origin of replication as disclosed therein. In one preferred embodiment, a nucleic acid expression construct comprises an expression control sequence operably linked to a polynucleotide encoding a hybrid polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, or 16. In another preferred embodiment, the nucleic acid expression construct has an inducible promoter, which may be *lac, tac, trc, ara, trp*, λ phage, T7 phage, and T5 phage promoter, and more preferably is a T5 phage promoter/*lac* operator expression control sequence as set forth in SEQ ID NO:17. The "expression control sequence" refers to any sequences sufficient to allow expression of a protein of interest in a host cell, including one or more promoter sequences, enhancer sequences, operator sequences (*e.g., lacO*), and the like. In certain embodiments, the hybrid polypeptide-encoding nucleic acid is in a plasmid, more preferably in plasmid pT5 (SEQ ID NO:17) and the host cell is a bacterium, most preferably *Escherichia coli.*

It should be understood that hybrid polypeptide-encoding nucleic acid may be a variant of the natural sequence due to, for example, the degeneracy of the genetic code (including alleles). Briefly, such "variants" may result from natural polymorphisms or may be synthesized by recombinant methodology (e.g., to obtain codon optimization for expression in a particular host) or chemical synthesis, and may differ from wild-type polypeptides by one or more amino acid substitutions, insertions, deletions, or the like.

Variants encompassing conservative amino acid substitutions include, for example, substitutions of one aliphatic amino acid for another, such as Ile, Val, Leu, or Ala or substitutions of one polar residue for another, such as between Lys and Arg, Glu and Asp, or Gln and Asn. Such substitutions are well known in the art to provide variants having similar physical properties and functional activities, such as for example, the ability to elicit and cross-react with similar antibodies. Other variants include nucleic acids sequences that encode a hybrid polypeptide having at least 50%, 60%, 70%, 80%, 90% or 95% amino acid identity to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16. Preferred embodiments are those having greater than 90% or 95% identity with the amino acid sequence of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16. As will be appreciated by those of ordinary skill in the art, a nucleotide sequence encoding an hybrid polypeptide or variant thereof may differ from the native sequences presented herein due to codon degeneracy, nucleotide polymorphism, or nucleotide substitution, deletion or insertion. Thus, in certain aspects the present invention includes all degenerate nucleic acid molecules that encode polypeptides and peptides comprising the amino acid sequence of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16. In another aspect, included are nucleic acid molecules that encode hybrid polypeptide variants having conservative amino acid substitutions or deletions or substitutions such that the hybrid polypeptide variant retains at least one epitope capable of eliciting antibodies specific for one or more streptococcal serotypes.

In certain aspects, a nucleic acid sequence may be modified to encode a hybrid polypeptide variant wherein specific codons of the nucleic acid sequence have been changed to codons that are favored by a particular host and can result in enhanced levels of expression (*see, e.g.*, Haas *et al.*, *Curr. Biol. 6*:315, 1996; Yang *et al.*, *Nucleic Acids Res. 24*:4592, 1996). For example, certain codons of the immunogenic peptides obtained from streptococcal M proteins were optimized, without changing the primary sequence of the peptides, for improved expression in *Escherichia coli*. By way of illustration and not limitation, eleven of the thirteen arginine (Arg) codons of AGG/AGA in the hexavalent hybrid polypeptide as set forth in SEQ ID NO:9 were changed to the Arg codons of CGT/CGC as set forth in SEQ ID NO:1. Similarly, twelve of twenty AGG/AGA Arg

codons of SEQ ID NO:15 were optimized to CGT/CGC codons as set forth in SEQ ID NO:8; seven of thirteen AGG/AGA Arg codons of SEQ ID NO:11 were optimized to CGT/CGC codons as set forth in SEQ ID NO:3; and five of nine AGG/AGA Arg codons of SEQ ID NO:13 were optimized to CGT/CGC codons as set forth in SEQ ID NO:5. As is known in the art, codons may be optimized for whichever host the hybrid polypeptide is to be expressed in, including without limitation bacteria, fungi, insect cells, plant cells, and mammalian cells. Additionally, codons encoding different amino acids may be changed as well, wherein one or more codons encoding different amino acids may be altered simultaneously as would best suit a particular host (*e.g.*, codons for arginine, glycine, leucine, and serine may all be optimized or any combination thereof). Alternatively, codon optimization may result in one or more changes in the primary amino acid sequence, such as a conservative amino acid substitution, addition, deletion, or combination thereof.

While particular embodiments of isolated nucleic acids encoding hybrid polypeptides are depicted in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15, within the context of the present invention, reference to one or more isolated nucleic acids includes variants of these sequences that are substantially similar in that they encode native or non-native hybrid polypeptides with similar structure and ability to elicit serospecific antibodies to at least one immunogenic peptide subunit contained in the hybrid polypeptides of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16. As used herein, the nucleotide sequence is deemed to be "substantially similar" if: (a) the nucleotide sequence is derived from the coding region of a *emm* gene isolated from a streptococcus (including, for example, portions of the sequence or allelic variations of the sequences discussed above) and contains a M protein epitope with substantially the same ability to elicit opsonic antibodies protective against streptococci that are not tissue cross-reactive; (b) the nucleotide sequence is capable of hybridization to the nucleotide sequences of the present invention under moderate or high stringency; (c) the nucleotide sequences are degenerate (*i.e.*, sequences which code for the same amino acids using a different codon sequences) as a result of the genetic code to the nucleotide sequences defined in (a) or (b); or (d) is a complement of any of the sequences described in (a), (b) or (c).

As used herein, two nucleotide sequences are said to "hybridize" under conditions of a specified stringency when stable hybrids are formed between substantially complementary nucleic acid sequences. Stringency of hybridization refers to a description of the environment under which hybrids are annealed and washed, which typically includes ionic strength and temperature. Other factors that might affect hybridization include the probe size and the length of time the hybrids are allowed to form. For example, "high," "medium" and "low" stringency encompass the following conditions or equivalent conditions thereto: high stringency is 0.1 x SSPE or SSC, 0.1% SDS, 65°C; medium stringency is 0.2 x SSPE or SSC, 0.1% SDS, 50°C; and low stringency is 1.0 x SSPE or SSC, 0.1% SDS, 50°C. As used herein, the term "high stringency conditions" means that one or more sequences will remain hybridized only if there is at least 95%, and preferably at least 97%, identity between the sequences. In preferred embodiments, the nucleic acid sequences that remain hybridized to a hybrid polypeptide-encoding nucleic acid molecule encode polypeptides that retain at least one epitope of a hybrid polypeptide encoded by a nucleic acid of any one of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15.

Methods for producing the hybrid polypeptides of the subject invention are provided as well wherein any of the nucleic acid molecules and host cells described herein may be used. In a preferred embodiment, a method of producing a hybrid polypeptide comprises culturing a host cell containing a nucleic acid expression vector comprising at least one expression control sequence operably linked to a nucleic acid molecule encoding a hybrid polypeptide, such as a hybrid polypeptide as set forth in any one of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, or 16, under conditions and for a time sufficient for expression of the polypeptide. In one particularly preferred embodiment, a hybrid polypeptide is produced by this method, and more preferably the hybrid polypeptides produced are of SEQ ID NOS:10, 12, 14, or 16, and more preferably the hybrid polypeptides produced are of SEQ ID NOS:2, 4, 6, or 8.

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

GENERATION OF RECOMBINANT MULTIVALENT AND INDIVIDUAL STREPTOCOCCAL PROTEINS

Once the specific 5' sequences of each *emm* and *spa* gene had been selected for inclusion in the vaccine, they were used to design four hybrid nucleic acid molecules, each containing 6-7 *emm* and/or *spa* coding sequences linked in tandem by unique restriction enzyme recognition sites (Figures 1-5). The four hybrid nucleic acid molecules were constructed using PCR-generated *emm* or *spa* nucleic acid molecules that were amplified from streptococcal genomic DNA of the corresponding serotype using oligonucleotide forward and reverse primers containing restriction enzyme sites at the 5' end. The PCR-generated fragments were purified, digested with the appropriate restriction enzymes, ligated using methods previously described (Dale *et al.*, *J. Immunol. 151*:2188, 1993; Dale, *Vaccine 17*:193, 1999), and then sequentially cloned into the expression vector pT5. Plasmid pT5 (SEQ ID NO:17) comprises a bacteriophage T5 promoter operably to *lac* operators, which means expression from the T5 promoter can be induced with isopropyl-beta-D-thiogalactopyranoside (IPTG).

The 5' *emm* fragment was reiterated at the 3' end of each hybrid nucleic acid molecule based on observations that an amino-terminal M protein peptide reiterated on the carboxy-terminus of a hybrid polypeptide appears to enhance or protect the immunogenicity of the adjacent M protein peptides (*see* Dale, *Vaccine 17*:193, 1999; WO 99/13084). In addition, the *emm* nucleic acid molecule cloned at the 3'-end of the hybrid molecule was engineered to include at the 5' end the following: (a) at least six histidine codons, (b) a *Xho*I restriction enzyme site, (c) optionally one or more amino acid codons (*e.g.*, cysteine), and (d) at least one stop codon (TAA or TAG). Several of the Arg codons from each of the hexavalent and septavalent polypeptides were optimized (*see* Figures 2-5), without changing the primary amino acid sequence, for expression in *E. coli*, which yielded from about a 2-fold to about a 10-fold increase in polypeptide production.

The nucleic acid and amino acid sequences of the hybrid polypeptides used to make admixtures for immunizing rabbits and/or humans are set forth in SEQ ID NOS:1-16. Each expression plasmid construct of pT5 was used to transform *E. coli* strain JM105 (genotype F' *tra*D36 *lacI*^{q}Δ(*lacZ*)M15*proA*$^+$B$^+$/*thi rpsL* (str^r) *endA sbcB*15 *sbcC*? *hsdR*4($r_k^-m_k^-$) Δ(*lac-pro*AB). The sequence identity of each hybrid DNA molecule transformed into JM105 *E. coli* was verified by sequencing both strands. Expression of each component fusion protein was detected by SDS-PAGE analysis using whole cell lysates before and after 1 mM IPTG induction.

In addition to the hybrid multivalent polypeptides generated to immunize a particular subject, recombinant homodimeric peptides comprising a single serotype immunogenic peptide were expressed and purified to test immune sera elicited by the hybrid streptococcal polypeptides. Each *emm* or *spa* gene fragment, included in the hybrid nucleic acid molecules described above, was independently amplified by PCR, purified, and cloned sequentially into the expression vector pT5 as an in-frame dimer with a restriction enzyme site between each coding sequence. Each PCR-generated sequence was verified by sequencing both strands of the dimer-encoding nucleic acid molecule. Expression of each peptide in transformed JM105 *E. coli* was detected by SDS-PAGE analysis, as described above.

Each hybrid polypeptide encoded by the hybrid nucleic acid molecules was further analyzed by BlastP (matrix BLOSUM62) to assure that there were no significant homologies with human proteins in the GenBank database. The linking restriction enzyme sites between the *emm* and/or *spa* encoding nucleic acid molecules were selected to avoid creating a sequence encoding potential human tissue cross-reactive epitopes. Specifically, the two amino acid residues encoded by each restriction enzyme site along with the six flanking M protein or Spa residues on each side of the sites (14 residues in total) were searched using BlastP (matrix PAM30) to detect potential homologies with human proteins in the GenBank database. Any restriction enzyme site matching more than four contiguous amino acids of a human protein sequences was not used.

EXAMPLE 2

PURIFICATION OF MULTIVALENT AND INDIVIDUAL STREPTOCOCCAL PROTEINS

Each hybrid polypeptide and individual dimeric peptide was purified separately. Cell paste of *E. coli* JM105 expressing His-tagged hybrid polypeptide was lysed in phosphate buffered saline (PBS) by microfluidation (Microfluidics, Inc., Newton, MA). After centrifugation, the clarified lysate was batch adsorbed to nickel-loaded affinity chelate resin (Tosoh Biosep, Montgomeryville, PA), washed and eluted with a step gradient of imidazole in PBS. Fractions containing the eluted hybrid polypeptide were pooled, pumped through a 5 ml HITRAP® Q anion exchange cartridge (Amersham Pharmacia Biotech, Piscataway, NJ) and concentrated in a stir cell (Millipore, Bedford, MA). The concentrated hybrid polypeptide was further purified by size exclusion chromatography using a Superdex 200 column (600 cm length, Amersham Pharmacia Biotech) equilibrated with PBS. Fraction purity was monitored using SDS-PAGE and fractions containing pure hybrid polypeptide were pooled and stored at –20°C until use. Purity, identity and concentration of the hybrid polypeptide were further assessed by reverse-phase HPLC, amino acid analysis, and electrospray mass spectrometry.

Individual dimeric peptides were similarly purified, but with notable differences. *E. coli* JM105 expressing 6xHis-tagged dimeric peptide were lysed in PBS containing 8M urea for 3 hours with stirring. After centrifugation, the clarified lysate was batch adsorbed to nickel-loaded affinity chelate resin (Tosoh Biosep) and washed and eluted with a step gradient of imidazole in PBS. The eluted dimeric peptide was then loaded onto a preparative reverse-phase C4 column (Vydac, Hesperia, CA), washed, and eluted with increasing concentrations of acetonitrile in water containing 0.1% trifluoroacetic acid. Fractions of eluted dimeric peptide were monitored using SDS-PAGE. Fractions containing purified dimeric peptide were pooled and dialyzed against PBS before storage at –20°C.

EXAMPLE 3

FORMULATION OF HYBRID POLYPEPTIDE COCKTAIL AND IMMUNIZATION OF RABBITS

The four multivalent polypeptides (Hexa A.1 [SEQ ID NO:10], Septa B.2 [SEQ ID NO:16], Septa C.2 [SEQ ID NO:4], and Septa D.1 [SEQ ID NO:14]) were mixed in equimolar amounts and adsorbed to alum (REHYDRAGEL®, low viscosity, Reheis, Inc., Berkeley Heights, NJ) to achieve a final protein concentration of 800 μg/ml and a final alum concentration of 1.5 μg/ml. This cocktail of immunogenic polypeptides represent at least 27 antigens.

New Zealand white rabbits were each immunized with 400 μg (*i.e.*, about 100 μg of each multivalent polypeptide) of the vaccine via the intramuscular route at either 0, 4, and 8 weeks or 0, 4, and 16 weeks (*see* Dale, *Vaccine 17*:193, 1999). Serum was obtained prior to the first injection and two weeks after the final injection.

EXAMPLE 4

ELISA USING SERUM FROM IMMUNIZED RABBITS

Type-specific antibodies were detected by enzyme-linked immunosorbent assays (ELISAs) essentially by methods previously described (McLellan *et al.*, *Infect. Immun. 69*:2943, 2001). Briefly, microtiter wells were coated with purified recombinant dimeric M peptides (*i.e.*, copying the vaccine subunits and used as solid-phase antigens). Wells without peptide but containing all other reagents served as negative controls. The ELISAs were performed using pre-immune and immune rabbit sera. The sera were serially diluted in PBS (pH 7.4) with 0.05% TWEEN® 20, added to the wells, and incubated at 37°C for 2 hours. The wells were washed with 0.15% saline-TWEEN® 20. A horse radish peroxidase-conjugated goat immunoglobulin G (IgG) to rabbit immunoglobulins (IgG, IgA, and IgM) (ICN Biomedicals, Aurora, OH) diluted 1:2,000 was added and incubated at 37°C for 2 hours. The wells were then washed, 5-aminosalicylic acid was added, and the

A_{450} was recorded after 15 minutes in an MR 600 microplate reader (Dynatech Laboratories, Inc., Chantilly, VA).

The immune sera from rabbits (obtained 2 weeks after the final immunization) contained high titers of antibodies to the majority of the M peptides contained in the vaccine (Figure 7). Antibody titers were determined for each of the 26 M peptides and Spa, a new protective antigen of group A streptococci (Dale *et al., J. Clin. Invest. 130*:1261, 1999). All pre-immune titers were less than 200. Out of the 81 immune serum titers determined (27 antigens x 3 rabbits), 69 titers (85%) increased by four-fold or greater over the pre-immune levels (Figure 7).

EXAMPLE 5

OPSONIZATION ASSAYS USING SERUM FROM IMMUNIZED RABBITS

Opsonic antibodies were detected by *in vitro* opsonization assays, essentially as previously described (Beachey *et al., J. Exp. Med. 145*:1469, 1977). The test mixture consisted of 0.05 ml of a standard suspension of streptococci grown to mid-log phase, 0.05 ml test serum, and 0.2 ml whole, heparinized (10 U/ml) nonimmune human blood. For these assays, the number of streptococcal CFU per leukocyte was approximately 10. The tubes were rotated end-over-end for 45 minutes at 37°C. Smears were then made on glass slides and stained with Wright's stain (Sigma Diagnostics, St. Louis, MO). Opsonization was quantitated by counting 50 consecutive neutrophils and calculating the percentage with associated streptococci (percent opsonization).

The pre-immune sera from all three rabbits resulted in ≤10% opsonization of each of the 26 serotypes tested (data not shown), indicating that the donor blood used for these assays did not contain antibodies against the test organism and that each organism was fully resistant to opsonization in nonimmune blood. Using 30% opsonization in the presence of immune serum as a positive threshold result (*i.e.*, three or more times the pre-immune level), 18 of the 26 serotypes (69%) were opsonized by at least one of three immune rabbit sera (Figure 8).

EXAMPLE 6

BACTERICIDAL ASSAYS USING SERUM FROM IMMUNIZED RABBITS

Bactericidal assays were performed similar to Example 5 (Lancefield, *J. Exp. Med. 106*:525, 1957) except that 0.05ml of Todd-Hewitt broth containing fewer bacteria was added to 0.1 ml of test serum and 0.35 ml of blood and the mixture was rotated for three hours at 37°C. Then 0.1 ml aliquots of this mixture were added to melted sheep's blood agar, pour plates were prepared, and viable organisms (CFU) were counted after overnight incubation at 37°C. For each serotype tested, three different inocula were used to assure that the growth in blood containing pre-immune serum was optimal and was quantifiable. The results are expressed as percent killing, which was calculated using the following formula: [(CFU after three hours growth with pre-immune serum) – (CFU after three hours growth with immune serum)] ÷ [CFU after three hours growth with pre-immune serum] x 100. Only those assays that resulted in growth of the test strain to at least eight generations in the presence of pre-immune serum were used to express percent killing in the presence of immune serum.

In all experiments, the test mixture containing pre-immune serum resulted in growth of the organisms to eight generations or more (data not shown), again indicating that the human blood did not contain opsonic antibodies against the test strains and that each organism was fully resistant to bactericidal killing in nonimmune blood. Using 50% reduction in growth after the three-hour rotation in immune serum compared to the pre-immune serum (percent killing), bactericidal activity was observed against 22 of the 26 serotypes tested (Figure 9). When the results of the opsonization and bactericidal assays were combined, 24 of the 26 serotypes (92%) tested were opsonized by the immune sera in one or both assays.

EXAMPLE 7

ASSAYS TO DETECT TISSUE CROSS-REACTIVE ANTIBODIES IN IMMUNIZED RABBITS

Rabbit immune sera raised against a composition comprising a cocktail of four different hybrid polypeptides (*i.e.*, a 27-valent vaccine) were tested for the presence of tissue cross-reactive antibodies by indirect immunofluorescence assays (Dale and Beachey, *J. Exp. Med. 161*:113, 1985) using frozen sections (4 μm) of human myocardium, kidney, basal ganglia, cerebral cortex, and cartilage. The sections were placed on gelatin-coated slides and fixed with 1% paraformaldehyde for 10 min. The slides were washed with PBS, incubated with the immune sera diluted 1:5 in PBS for 30 minutes at room temperature, and washed thoroughly in PBS. The sections were then incubated with fluorescein-conjugated goat anti-rabbit IgG (Cappel, West Chester, PA) at a dilution of 1:40 in PBS for 30 minutes at room temperature. After washing, the slides were mounted in Gelvatol and examined in a fluorescent microscope. Rabbit anti-sera known to cross-react with human myocardium, kidney, and brain were used as positive controls and rabbit pre-immune sera as negative controls.

While the hybrid polypeptides elicited opsonic antibodies to most GrAS serotypes tested, none of the hybrid polypeptides elicited human cross-reactive antibodies. This indicates that the hybrid polypeptides, alone or in combination, do not contain potentially harmful autoimmune epitopes.

EXAMPLE 8

BACTERICIDAL ACTIVITY OF SERUM FROM IMMUNIZED RABBITS AGAINST OTHER SEROTYPES

Type 4 streptococci are relatively common causes of uncomplicated pharyngitis and invasive infections. Type 4 organisms currently account for 3.8% of all invasive infections and 8.6% of all pharyngitis isolates in the ongoing U.S. surveillance program (personal communication, S.T. Shulman). Not wishing to be bound by theory, it appears that purified recombinant type 4 M protein either does not evoke opsonic

antibodies or the type 4 streptococcal strains are resistant to opsonization. For this reason, the *emm*4 gene fragment was not included in the composition comprising the four hybrid polypeptides of Hexa A.1 [SEQ ID NO:10], Septa B.2 [SEQ ID NO:16], Septa C.2 [SEQ ID NO:4], and Septa D.1 [SEQ ID NO:14] (*i.e.*, the 27-valent vaccine). To determine whether any of the antibodies elicited by the 27-valent vaccine might be directed against cross-reactive opsonic epitopes on the surface of type 4 streptococci, bactericidal assays were performed using seven clinical isolates obtained from the U.S. Streptococcal Pharyngitis Surveillance Program (Figure 10).

Interestingly, bactericidal activity was detected against five of the seven strains of type 4 streptococci. The strains isolated from patients in Florida (FL9) and Illinois (IL23 and IL12) were opsonized by both of the 27-valent antisera tested. One of the two antisera opsonized both isolates from California (CA8 and CA12), while the strains from Connecticut (CT5) and South Dakota (SD32) were not opsonized by either antiserum (Figure 10). The results suggest that the 27-valent cocktail vaccine can evoke antibodies that cross-react with protective epitopes on the surface of some strains of type 4 streptococci. In addition, the data indicate that type 4 streptococci may be a heterogeneous group of organisms that express different protective epitopes even though they all express the type-specific M4 protein.

EXAMPLE 9

FORMULATION OF HYBRID POLYPEPTIDE COCKTAIL AND IMMUNIZATION OF HUMANS

The formulated bulk vaccine consists of the four recombinant proteins (Hexa A.3 [SEQ ID NO:2], Septa B.3a [SEQ ID NO:8], Septa C.2 [SEQ ID NO:4], and Septa D.3 [SEQ ID NO:6]) adsorbed onto aluminum hydroxide (final alum concentration of about 1.5 μg/ml) and diluted to a target concentration of about 400 μg/ml or about 800 μg/ml (100 μg/ml or 200 μg/ml, respectively, of each hybrid polypeptide) with phosphate buffered saline. This cocktail of immunogenic polypeptides represents at least 27 antigens.

Briefly, the calculated volumes of the purified hybrid polypeptides in PBS are measured out and added to a sterile polystyrene media bottle. The mixture is stirred until homogenous, and then diluted with an equal volume of sterile water for injection. This dilution step reduces the concentration of Na_2HPO_4 and NaCl in the mixture to the desired final concentration (5 mM phosphate, 150 mM NaCl, pH 7.5). The pooled HYBRID polypeptides are then passed through a sterilizing filter unit (MILLIPAK® 20, Millipore, Bedford, MA) using a peristaltic pump. If some peptides require some special conditions to homogenize or dilute, then the solutions are made separately and subsequently mixed with an adjustment to pH, as needed.

The required volume of recombinant hybrid polypeptides is measured out, and added to a formulation bottle. REHYDRAGEL® (low viscosity, Reheis, Inc., Berkeley Heights, NJ) is received as a sterile suspension of aluminum hydroxide in water for injection and is used without further preparation. The required volume of REHYDRAGEL® is measured out and added to the formulation bottle while stirring. The pH of the mixture is measured and adjusted to pH 7.5 - 7.7 using 1 M NaOH. Finally, formulation buffer is added to achieve the correct final volume and the mixture stirred for an additional 16 - 20 hours at room temperature. The bulk vaccine is divided into containers (sterile), samples are taken for testing, and the formulated vaccine is stored at 2-8 °C.

Human volunteers were screened and 30 healthy subjects aged 18-50 years were enrolled. Each subject was immunized with 400 μg of the cocktail vaccine composition via the intramuscular route at 0, 30, and 120 days. Serum was obtained prior to the first injection and at days 14, 30, 44, 60, 120, 134 and 150.

EXAMPLE 10

ELISA Using Serum from Immunized Human Subjects

Type-specific antibodies were detected by ELISA, similar to the methods described in Example 4 (*see also* McLellan *et al.*, *Infect. Immun.* *69*:2943, 2001). Briefly,

microtiter wells were coated with purified recombinant dimeric M peptides (*i.e.*, copying the vaccine subunits and used as solid-phase antigens). Wells without human sera but containing all other reagents served as negative controls. The ELISAs were performed using the collected human sera. The sera were serially diluted in PBS (pH 7.4) with 0.1% BSA and 0.05% TWEEN® 20, added to the wells, and incubated at 37°C for 2 hours. The wells were washed with PBS-0.05%TWEEN® 20. A horseradish peroxidase-conjugated goat immunoglobulin G (IgG) to human immunoglobulins (IgG, IgA, and IgM) (Jackson ImmunoResearch Laboratories, West Grove, PA) diluted 1:2,000 was added and incubated at 37°C for 1 hour. The wells were then washed, 1-Step™ Turbo TMB-ELISA substrate (Pierce, Rockford, IL) was added, 1 N sulfuric acid was added after 30 minutes. The $A_{450/595}$ was recorded in a V_{MAX}® microplate reader (Molecular Devices, Sunnyvale, CA).

All of the immune sera from the human subjects (obtained at Day 134, which was two weeks after the third injection on Day 120) showed a statistically significant rise in antibody titer for all M antigens and the Spa antigen, with a p value < 0.001 (Figure 11). All subjects had a baseline level pre-immune titer for some antigens (Day 0) as seen in Figure 11 (Day 0, open bars). The geometric mean fold increase in serum antibody to all of the antigens in the 27-valent cocktail was 12.6-fold, with a minimum increase of three-fold and a maximum increase of over 68-fold. Overall, 26 out of the 27 antigens represented in the vaccine evoked at least a four-fold mean increase in antibody titer.

EXAMPLE 11

BACTERICIDAL ASSAYS USING SERUM FROM IMMUNIZED HUMAN SUBJECTS

Bactericidal assays were performed essentially as described in Example 6 (*see also* Lancefield, *J. Exp. Med. 106*:525, 1957). Only those assays that resulted in growth of the test strain in the presence of Day 0 serum were used to express percent killing in the presence of Day 134 serum.

All subjects that showed a rise in antibody titers over baseline level pre-immune titers also showed an increase over baseline levels bactericidal activity

(Figure 12). The geometric mean antibody titer increase was determined as in Example 10 and converted to log normal (Ln) for comparison to functional activity (bactericidal killing) of the same immune sera. Thus, there is a quantitative correlation between increased antibody titers and the ability of antibodies to induce bacterial killing.

EXAMPLE 12

ASSAYS TO DETECT TISSUE CROSS-REACTIVE ANTIBODIES IN SERUM FROM IMMUNIZED HUMAN SUBJECTS

Sera collected from human subjects, who were immunized with a 27-valent vaccine composition comprising a cocktail of four different hybrid polypeptides (Hexa A.3 [SEQ ID NO:2], Septa B.3a [SEQ ID NO:8], Septa C.2 [SEQ ID NO:4], and Septa D.3 [SEQ ID NO:6]), were tested for the presence of tissue cross-reactive antibodies as essentially described in Example 7.

Similar to the results observed in rabbits, none of the hybrid polypeptides elicited human cross-reactive antibodies. This indicates that the hybrid polypeptides, alone or in combination, do not contain potentially harmful human autoimmune epitopes.

EXAMPLE 13

COMPARISON OF HEXAVALENT VACCINE VERSUS 27-VALENT COCKTAIL VACCINE USED TO IMMUNIZED HUMAN SUBJECTS

A similar human trial was performed using single hexavalent polypeptide Hexa 1.2, which has a structure of M24-M5-M6-M19-M1-M3-M24 (*see* Dale, *Vaccine 17*:193, 1999; WO 99/13084). Human volunteers were screened and 11 healthy subjects were each immunized three times with 100 μg of Hexa 1.2 in the same formulation buffer as used for the 27-valent vaccine composition comprising a cocktail of four different hybrid polypeptides (Hexa A.3 [SEQ ID NO:2], Septa B.3a [SEQ ID NO:8], Septa C.2 [SEQ ID NO:4], and Septa D.3 [SEQ ID NO:6]). All six of the group A streptococcal

antigens of Hexa 1.2 (M24, M5, M6, M19, M1, and M3) are represented in two of the four multivalent polypeptides that comprise the 27-valent vaccine. ELISAs were then performed on sera from immunized subjects to identify type-specific antibodies for each of the Hexa 1.2 antigens, essentially as described in Example 10. The geometric mean titers were calculated using antibody titers from the 11 subjects that received the Hexa 1.2 multivalent polypeptide and using antibody titers from the 30 subjects in the 27-valent cocktail of four different multivalent polypeptides clinical trial, respectively (Figure 13). The fold-increase in antibody titers represents the fold rise in geometric mean titers after immunization as compared to the geometric mean titers before immunization.

Surprisingly, the subjects immunized with the 27-valent composition showed a much greater fold-increase in antibody titers overall than did the subjects who received the hexavalent composition (Figure 13). The 27-valent showed a greater fold-increase for all M antigens, ranging from about 1.1- to about a 4-fold increase. As noted above, the increase in antibody titer also correlates with an increase in bactericidal activity. Hence, the four multivalent peptides together unexpectedly showed a synergistic effect in evoking an immune response as compared to a single hexavalent peptide.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

1. A hybrid polypeptide, comprising at least six different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 30 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least M5, M6, M14, M19, M24, and M29.

2. The hybrid polypeptide according to claim 1 wherein the amino-terminal immunogenic peptide of the polypeptide is M24.

3. The hybrid polypeptide according to claim 2 wherein the polypeptide is recombinant and the immunogenic peptides are linked in tandem, the polypeptide having a structure of M24-M5-M6-M19-M29-M14-M24.

4. A hybrid polypeptide according to claim 3 wherein the polypeptide comprises the amino acid sequence of SEQ ID NO:2.

5. A hybrid polypeptide, comprising at least seven different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 35 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least M2, M11, M22, M33, M43, M59, and M94.

6. The hybrid polypeptide according to claim 5 wherein the amino-terminal immunogenic peptide of the polypeptide is M2.

7. The hybrid polypeptide according to claim 6 wherein the polypeptide is recombinant and the immunogenic peptides are linked in tandem, the polypeptide having a structure of M2-M43-M94-M22-M11-M59-M33-M2.

8. A hybrid polypeptide according to claim 7 wherein the polypeptide comprises the amino acid sequence of SEQ ID NO:4.

9. A hybrid polypeptide, comprising at least seven different linked immunogenic peptides, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 40 amino acids, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least M75, M76, M77, M89, M92, M101, and M114.

10. The hybrid polypeptide according to claim 9 wherein the amino-terminal immunogenic peptide of the polypeptide is M89.

11. The hybrid polypeptide according to claim 10 wherein the polypeptide is recombinant and the immunogenic peptides are linked in tandem, the polypeptide having a structure of M89-M101-M77-M114-M75-M76-M92-M89.

12. A hybrid polypeptide according to claim 11 wherein the polypeptide comprises the amino acid sequence of SEQ ID NO:6.

13. A hybrid polypeptide, comprising at least seven different immunogenic peptides linked in tandem, wherein each peptide comprises an amino-terminal portion of a streptococcal M protein of at least 50 amino acids and, and wherein the polypeptide has an amino-terminal peptide that is reiterated as a carboxy-terminal peptide and the polypeptide is

capable of eliciting an immune response against more than one antigen of group A streptococci comprising at least Spa, M1.0, M1.2, M3, M12, M18, and M28.

14. The hybrid polypeptide according to claim 13 wherein the amino-terminal immunogenic peptide of the polypeptide is M1.0.

15. The hybrid polypeptide according to claim 14 wherein the polypeptide is recombinant and the immunogenic peptides are linked in tandem, the polypeptide having a structure of M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0.

16. A hybrid polypeptide according to claim 15 wherein the polypeptide comprises the amino acid sequence of SEQ ID NO:8.

17. The hybrid polypeptide according to any one of claims 3, 7, 11, and 15 wherein the immunogenic peptides are linked by at least two amino acids encoded by a nucleic acid sequence that is a restriction enzyme recognition site.

18. The hybrid polypeptide according to claim 17 wherein the nucleic acid sequence is a restriction enzyme recognition site; the recognition site being at least one of *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Nco*I, *Nhe*I, *Pml*I, *Pst*I, *Sal*I, and *Xho*I.

19. The hybrid polypeptide according to any one of claims 1-16 wherein the polypeptide is capable of eliciting at least one opsonic antibody that is not a tissue cross-reactive antibody in a subject.

20. The hybrid polypeptide according to claim 19 wherein the subject is a human or an animal.

21. The hybrid polypeptide according to any one of claims 1-16 further comprising a carboxy-terminal tag.

22. The hybrid polypeptide according to claim 21 wherein the carboxy-terminal tag is selected from the group consisting of alkaline phosphatase, β-galactosidase, hexahistidine, epitope tag DYKDDDDK (SEQ ID NO:18), epitope tag DLYDDDDK (SEQ ID NO:19), and GST.

23. The hybrid polypeptide according to claim 22 wherein the carboxy-terminal tag is hexahistidine.

24. The hybrid polypeptide according to any one of claims 1-16 further comprising at least one additional carboxy-terminal amino acid.

25. The hybrid polypeptide according to claim 24 wherein the additional carboxy-terminal amino acid is a D-amino acid.

26. The hybrid polypeptide according to claim 24 wherein the additional carboxy-terminal amino acid is cysteine.

27. The hybrid polypeptide according to claim 23 further comprising at least one additional carboxy-terminal amino acid.

28. The hybrid polypeptide according to claim 27 wherein the additional carboxy-terminal amino acid is cysteine.

29. A hybrid polypeptide comprising the amino acid sequence of SEQ ID NOS:2, 4, 6, or 8.

30. A nucleic acid molecule comprising a sequence encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8.

31. A nucleic acid expression construct comprising an expression control sequence operably linked to a polynucleotide encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8.

32. The construct according to claim 31 wherein the expression control sequence is an inducible promoter.

33. The construct according to claim 32 wherein the inducible promoter is selected from the group consisting of a *lac, tac, trc, ara, trp*, λ phage, T7 phage, and T5 phage promoter.

34. The construct according to claim 32 wherein the inducible promoter is a T5 phage promoter.

35. The construct according to claim 31 wherein the construct comprises a nucleic acid expression vector selected from the group comprising a plasmid, phagemid, shuttle vector, cosmid, and virus.

36. The construct according to claim 31 wherein the construct comprises a nucleic acid expression vector, wherein the vector is a plasmid.

37. The construct according to claim 36 wherein the plasmid is pT5 (SEQ ID NO:17).

38. A host cell containing a construct according to any one of claims 31-37.

39. The host cell of claim 38 wherein the host cell is selected from the group consisting of a bacterium, a yeast cell, a nematode cell, an insect cell, and a mammalian cell.

40. The host cell of claim 39 wherein the host cell is a bacterium, wherein the bacterium is *Escherichia coli*.

41. A method of producing a hybrid polypeptide, comprising culturing a host cell containing a nucleic acid expression vector comprising at least one expression control sequence operably linked to a nucleic acid molecule encoding a hybrid polypeptide of SEQ ID NOS:2, 4, 6, or 8, under conditions and for a time sufficient for expression of the polypeptide.

42. The method of claim 41 wherein the expression control sequence is an inducible promoter.

43. The method according to claim 42 wherein the inducible promoter is selected from the group consisting of a *lac, tac, trc, ara, trp*, λ phage, T7 phage, T5 phage promoter.

44. The method according to claim 42 wherein the inducible promoter is a T5 phage promoter.

45. The method according to claim 42 wherein the nucleic acid expression vector is pT5 (SEQ ID NO:17).

46. A hybrid polypeptide produced according to the method of claim 45.

47. A composition, comprising a pharmaceutically acceptable carrier and a hybrid polypeptide according to any one of claims 1-16, 29, and 46.

48. A composition, comprising a pharmaceutically acceptable carrier and a mixture of at least two of the hybrid polypeptides according to claims 1, 5, 9, and 13.

49. The composition according to claim 48 wherein the polypeptides comprise immunogenic peptides linked in tandem, wherein said at least two polypeptides comprise a structure of M24-M5-M6-M19-M29-M14-M24, M2-M43-M94-M22-M11-M59-M33-M2, M89-M101-M77-M114-M75-M76-M92-M89, or M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0.

50. A composition, comprising a pharmaceutically acceptable carrier and a mixture of the hybrid polypeptide according to claim 13 with at least one of the hybrid polypeptides according to any one of claims 1, 5, or 9.

51. The composition according to claim 50 wherein the polypeptides comprise immunogenic peptides linked in tandem, wherein the polypeptide according to claim 13 comprises a structure of M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0 and said at least one other polypeptide comprise a structure of M24-M5-M6-M19-M29-M14-M24, M2-M43-M94-M22-M11-M59-M33-M2, or M89-M101-M77-M114-M75-M76-M92-M89.

52. A composition, comprising a pharmaceutically acceptable carrier and a mixture of at least three of the hybrid polypeptides according to claims 1, 5, 9, and 13.

53. The composition according to claim 52 wherein the polypeptides comprise immunogenic peptides linked in tandem, wherein said at least three polypeptides comprise a structure of M24-M5-M6-M19-M29-M14-M24, M2-M43-M94-M22-M11-M59-M33-M2, M89-M101-M77-M114-M75-M76-M92-M89, or M1.0-M12-Spa-M28-M3-M1.2-M18-M1.0.

54. A composition, comprising a pharmaceutically acceptable carrier and a mixture of hybrid polypeptides according to claims 4, 8, 12, and 16.

55. The composition according to claim 47 wherein the polypeptides are in equimolar amounts.

56. The composition according to claim 47 wherein at least one of the hybrid polypeptides includes a Spa immunogenic peptide.

57. The composition according to claim 47 further comprising an adjuvant.

58. The composition according to claim 57 wherein the adjuvant is alum or Freund's.

59. The composition according to claim 54 further comprising an adjuvant, wherein said adjuvant is alum.

60. A method for preventing a microbial infection, comprising administering to a subject a composition according to claim 47 at a dose sufficient to elicit antibodies specific for one or more hybrid polypeptide, wherein the antibodies are opsonic and are not tissue cross-reactive.

61. The method for preventing an infection according to claim 60 wherein the microbial infection is a streptococcal infection.

62. The method for preventing an infection according to claim 60 wherein the streptococcal infection is a group A streptococcal infection.

63. The method for preventing an infection according to claim 60 wherein the hybrid polypeptide is administered by a route selected from the group consisting of enteral, parenteral, transdermal, transmucosal, and inhalation.

64. The method for preventing an infection according to claim 60 wherein the composition further comprises an adjuvant.

65. The method for preventing an infection according to claim 64 wherein the adjuvant is alum or Freund's.

66. The method for preventing an infection according to claim 60 wherein the subject is human or an animal.

67. A plurality of antibodies produced by the method according to claim 60.

68. The plurality of antibodies according to claim 67 wherein the antibodies further comprise at least one antibody specific for a M protein antigen not represented in a hybrid polypeptide.

69. The plurality of antibodies according to claim 68 wherein the M protein antigen not represented in a hybrid polypeptide is M4.

70. A method for treating or preventing a microbial infection, comprising administering to a subject a composition comprising a pharmaceutically acceptable carrier and a plurality of antibodies according to claims 67.

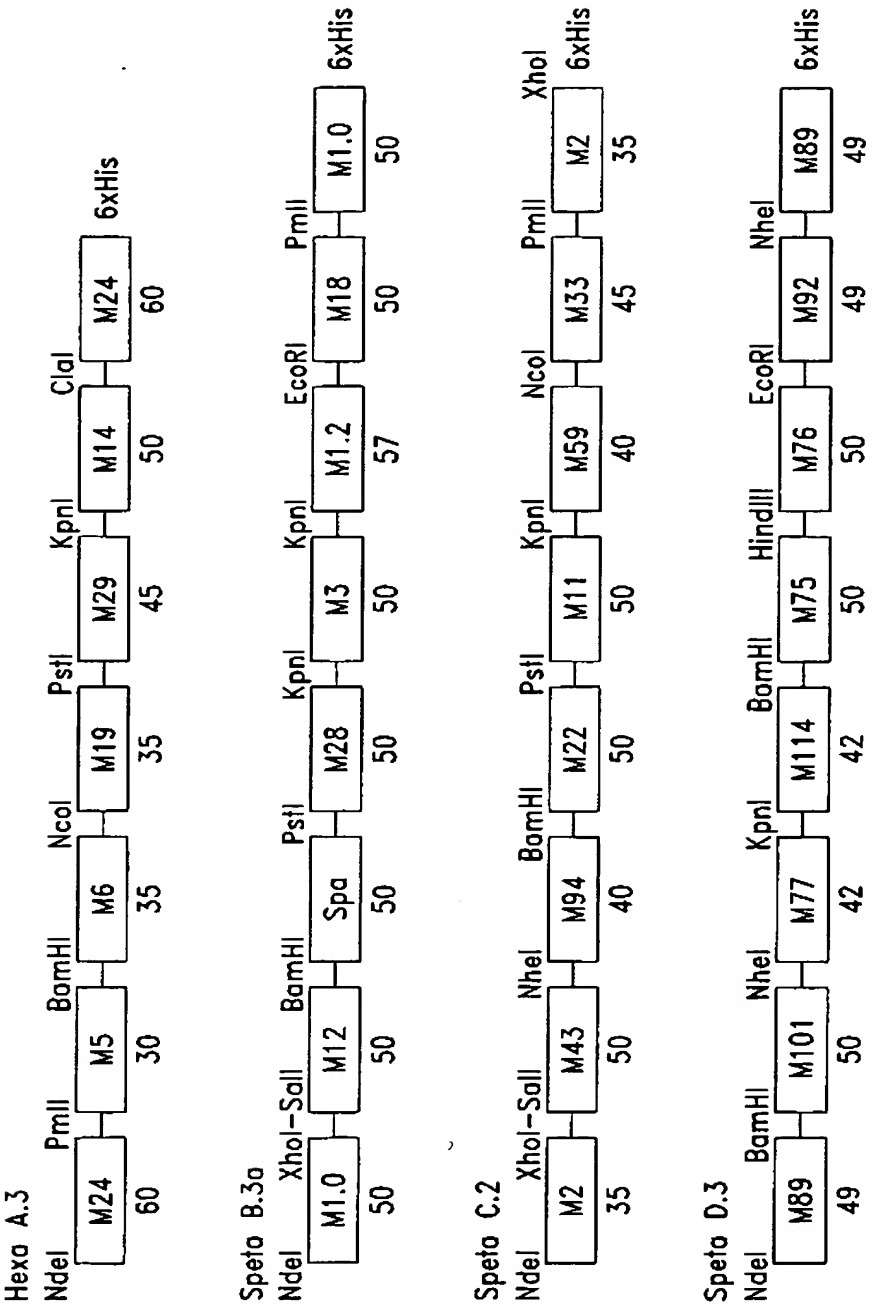

Hexa A.3
NdeI
M24
60
PmlI
M5
30
BamHI
M6
35
NcoI
M19
35
PstI
M29
45
KpnI
M14
50
ClaI
M24
60
6xHis
Speto B.3a
NdeI
M1.0
50
XhoI-SalI
M12
50
BamHI
Spa
50
PstI
M28
50
KpnI
M3
50
KpnI
M1.2
57
EcoRI
M18
50
PmlI
M1.0
50
6xHis
Speto C.2
NdeI
M2
35
XhoI-SalI
M43
50
NheI
M94
40
BamHI
M22
50
PstI
M11
50
KpnI
M59
40
NcoI
M33
45
PmlI
M2
35
XhoI
6xHis
Speto D.3
NdeI
M89
49
BamHI
M101
50
NheI
M77
42
KpnI
M114
42
BamHI
M75
50
HindIII
M76
50
EcoRI
M92
49
NheI
M89
49
6xHis

FIG. 1

Hexavalent A.3

```
        M24--->
             9          18          27          36          45          54
5'  ATG GTC GCG ACT CGC TCT CAG ACA GAT ACT CTG GAA AAA GTA CAA GAA CGT GCT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala

            63          72          81          90          99          108
    GAC AAG TTT GAG ATA GAA AAC AAT ACG TTA AAA CTT AAG AAT AGT GAC TTA AGT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser

           117         126         135         144         153         162
    TTT AAT AAT AAA GCG TTA AAA GAT CAT AAT GAT GAG TTA ACT GAA GAG TTG AGT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu Ser

                                PmlI    M5--->
           171         180         189         198         207         216
    AAT GCT AAA GAG AAA CTA CGT CAC GTG GCC GTG ACT CGC GGT ACA ATA AAT GAC
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Asn Ala Lys Glu Lys Leu Arg His Val Ala Val Thr Arg Gly Thr Ile Asn Asp

           225         234         243         252         261         270
    CCG CAA AGA GCA AAA GAA GCT CTT GAC AAG TAT GAG CTA GAA AAC CAT GAC TTA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Pro Gln Arg Ala Lys Glu Ala Leu Asp Lys Tyr Glu Leu Glu Asn His Asp Leu

                BamHI   M6--->
           279         288         297         306         315         324
    AAA ACT AAG GGA TCC CGT GTG TTT CCT CGC GGG ACG GTA GAA AAC CCG GAC AAA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Lys Thr Lys Gly Ser Arg Val Phe Pro Arg Gly Thr Val Glu Asn Pro Asp Lys

           333         342         351         360         369         378
    GCA CGA GAA CTT CTT AAC AAG TAT GAC GTA GAG AAC TCT ATG TTA CAA GCT AAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Ala Arg Glu Leu Leu Asn Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn
```

FIG. 2A

```
                NcoI    M19--->
        387         396         405         414         423         432
AAT GAC AAG TTA CCA TGG CGT GTG CGT TAT ACT CGC CAT ACG CCA GAA GAT AAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Asp Lys Leu Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys

        441         450         459         468         477         486
CTA AAA AAA ATT ATT GAC GAT CTT GAC GCA AAA GAA CAT GAA TTA CAA CAA CAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Lys Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln Gln Gln

                    PstI    M29--->
        495         504         513         522         531         540
AAT GAG AAG TTA TCT CTG CAG AAA GTG TAT ATT ACT CGT GGT ATG ACA AAA GAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Glu Lys Leu Ser Leu Gln Lys Val Tyr Ile Thr Arg Gly Met Thr Lys Glu

        549         558         567         576         585         594
GAC GTA GAA AAA ATT GCT AAC AAC CTT GAC ATA GAA AAC CAT GGG TTA AAA CAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asp Val Glu Lys Ile Ala Asn Asn Leu Asp Ile Glu Asn His Gly Leu Lys Gln

                                                            KpnI
        603         612         621         630         639         648
CAG AAT GAA CAG TTA TCT ACT GAT AAA CAA GGT CTT GAA GAA CAG AAT GGT ACC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Gln Asn Glu Gln Leu Ser Thr Asp Lys Gln Gly Leu Glu Glu Gln Asn Gly Thr

M14--->
        657         666         675         684         693         702
GAT CGC GTT AGT CGT TCT ATG TCA CGC GAT GAT CTA TTA AAC AGG GCT CAG GAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asp Arg Val Ser Arg Ser Met Ser Arg Asp Asp Leu Leu Asn Arg Ala Gln Asp

        711         720         729         738         747         756
CTT GAA GCA AAA AAC CAC GGG TTA GAA CAC CAG AAT ACT AAG TTA TCT ACT GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Glu Ala Lys Asn His Gly Leu Glu His Gln Asn Thr Lys Leu Ser Thr Glu
```

FIG. 2B

```
                                                                ClaI    M24--->
       765         774         783         792         801         810
AAT AAA ACG CTT CAA GAA CAA GCA GAA GCA CGC CAG AAA GAA ATC GAT GTC GCG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Lys Thr Leu Gln Glu Gln Ala Glu Ala Arg Gln Lys Glu Ile Asp Val Ala

       819         828         837         846         855         864
ACT CGC TCT CAG ACA GAT ACT CTG GAA AAA GTA CAA GAA CGT GCT GAC AAG TTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala Asp Lys Phe

       873         882         891         900         909         918
GAG ATA GAA AAC AAT ACG TTA AAA CTT AAG AAT AGT GAC TTA AGT TTT AAT AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser Phe Asn Asn

       927         936         945         954         963         972
AAA GCG TTA AAA GAT CAT AAT GAT GAG TTA ACT GAA GAG TTG AGT AAT GCT AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu Ser Asn Ala Lys

       981         990         999
GAG AAA CTA CGT CAC CAC CAC CAC CAC CAC TGA 3'
--- --- --- --- --- --- --- --- --- --- ---
Glu Lys Leu Arg His His His His His His stop
```

FIG. 2C

Septavalent B.3a

```
      M1.0--->
           9          18          27          36          45          54
5'  ATG AAC GGT GAT GGT AAT CCT AGG GAA GTT ATA GAA GAT CTT GCA GCA AAC AAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Met Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala Asn Asn

           63          72          81          90          99          108
    CCC GCA ATA CAA AAT ATA CGT TTA CGT CAC GAA AAC AAG GAC TTA AAA GCG AGA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp Leu Lys Ala Arg

                                                          (XhoI-SalI)M12--->
          117         126         135         144         153         162
    TTA GAG AAT GCA ATG GAA GTT GCA GGA AGA GAT TTT AAG AGA GCT CTC GAC GAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe Lys Arg Ala Leu Asp Asp

          171         180         189         198         207         216
    CAT AGT GAT TTA GTC GCA GAA AAA CAA CGT TTA GAA GAT TTA GGA CAA AAA TTT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    His Ser Asp Leu Val Ala Glu Lys Gln Arg Leu Glu Asp Leu Gly Gln Lys Phe

          225         234         243         252         261         270
    GAA AGA CTG AAA CAG CGT TCA GAA CTC TAC CTT CAG CAA TAC TAT GAT AAT AAA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Glu Arg Leu Lys Gln Arg Ser Glu Leu Tyr Leu Gln Gln Tyr Tyr Asp Asn Lys

                                                BamHI   Spa--->
          279         288         297         306         315         324
    TCA AAT GGA TAT AAA GGT GAC TGG TAT GTA CAA CAG TTA GGA TCC GAT TCA GTA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Ser Asn Gly Tyr Lys Gly Asp Trp Tyr Val Gln Gln Leu Gly Ser Asp Ser Val

          333         342         351         360         369         378
    AGT GGA TTA GAG GTG GCA GAC CCC TCT GAT AGT AAG AAA CTT ATT GAA TTA GGT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Ser Gly Leu Glu Val Ala Asp Pro Ser Asp Ser Lys Lys Leu Ile Glu Leu Gly
```

FIG. 3A

```
        387         396         405         414         423         432
TTG GCT AAA TAC CTT AAT GAT AAA TTA CCC TTT AAA ACT AAA GAA GAT TCA GAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Ala Lys Tyr Leu Asn Asp Lys Leu Pro Phe Lys Thr Lys Glu Asp Ser Glu

                                            PstI    M28--->
        441         450         459         468         477         486
ATT TTA TCA GAG TTA CGT GAT GTA TTA AAA AAT CTG CAG GAG TCT CCA AAA AGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ile Leu Ser Glu Leu Arg Asp Val Leu Lys Asn Leu Gln Glu Ser Pro Lys Ser

        495         504         513         522         531         540
ACT GAG ACT TCT GCT AAT GGA GCT GAT AAA TTA GCT GAT GCA TAC AAC ACA TTG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Glu Thr Ser Ala Asn Gly Ala Asp Lys Leu Ala Asp Ala Tyr Asn Thr Leu

        549         558         567         576         585         594
CTT ACT GAA CAT GAG AAA CTC AGA GAT GAG TAT TAT ACA TTA ATT GAT GCT AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Thr Glu His Glu Lys Leu Arg Asp Glu Tyr Tyr Thr Leu Ile Asp Ala Lys

                                    KpnI    M3--->
        603         612         621         630         639         648
GAA GAA GAA CCT CGC TAT AAA GCA TTG GGT ACC TTG TTA GAT CAG GTT ACA CAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Glu Glu Pro Arg Tyr Lys Ala Leu Gly Thr Leu Leu Asp Gln Val Thr Gln

        657         666         675         684         693         702
TTA TAT ACT AAA CAT AAT AGT AAT TAC CAA CAA TAT AAT GCA CAA GCT GGC AGA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Tyr Thr Lys His Asn Ser Asn Tyr Gln Gln Tyr Asn Ala Gln Ala Gly Arg

        711         720         729         738         747         756
CTT GAC CTG AGA CAA AAG GCT GAA TAT CTA AAA GGC CTT AAT GAT TGG GCT GAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Asp Leu Arg Gln Lys Ala Glu Tyr Leu Lys Gly Leu Asn Asp Trp Ala Glu

                            KpnI    M1.2--->
        765         774         783         792         801         810
CGC CTG TTA CAA GAG TTA AAT GGT ACC AAC AAT GAT GGT CGT TCT CGT GAC GTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Arg Leu Leu Gln Glu Leu Asn Gly Thr Asn Asn Asp Gly Arg Ser Arg Asp Val
```

FIG. 3B

```
        819         828         837         846         855         864
ACG GAA GAG ATT GCA GCA AAC AAT ACC ACA GTA CAA AAT ATA CGT TTA CGT AAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Glu Glu Ile Ala Ala Asn Asn Thr Thr Val Gln Asn Ile Arg Leu Arg Asn

        873         882         891         900         909         918
GAA AAC AAG AAC TTA AAA GCG AAA AAC GAG GAC TTA GAA GCG AGA TTA GAG AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Asn Lys Asn Leu Lys Ala Lys Asn Glu Asp Leu Glu Ala Arg Leu Glu Asn

                                                    EcoRI   M18--->
        927         936         945         954         963         972
GCA ATG AAT GTT GCA GGA CGC GAT TTT AAG CGT GCT GAA TTC GCA CCT CTT ACT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Met Asn Val Ala Gly Arg Asp Phe Lys Arg Ala Glu Phe Ala Pro Leu Thr

        981         990         999        1008        1017        1026
CGT GCT ACA GCA GAC AAT AAA GAC GAA TTA ATA AAA AGA GCT AAC GGT TAT GAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Arg Ala Thr Ala Asp Asn Lys Asp Glu Leu Ile Lys Arg Ala Asn Gly Tyr Glu

       1035        1044        1053        1062        1071        1080
ATA CAG AAC CAT CAG TTA ACA GTT GAG AAT AAA AAA TTA AAA ACT GAT AAG GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ile Gln Asn His Gln Leu Thr Val Glu Asn Lys Lys Leu Lys Thr Asp Lys Glu

                                        PmlI    M1.0--->
       1089        1098        1107        1116        1125        1134
CAG TTA ACA AAA GAG AAT GAT GAT TTA AAA CAC GTG AAC GGT GAT GGT AAT CCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Gln Leu Thr Lys Glu Asn Asp Asp Leu Lys His Val Asn Gly Asp Gly Asn Pro

       1143        1152        1161        1170        1179        1188
CGT GAA GTT ATA GAA GAT CTT GCA GCA AAC AAT CCC GCA ATA CAA AAT ATA CGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Arg Glu Val Ile Glu Asp Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg

       1197        1206        1215        1224        1233        1242
TTA CGT CAC GAA AAC AAG GAC TTA AAA GCG AGA TTA GAG AAT GCA ATG GAA GTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Arg His Glu Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val
```

FIG. 3C

```
      1251        1260        1269        1278        1287
GCA GGA CGT GAT TTT AAG CGT GCT CAC CAC CAC CAC CAC CAC TAA 3'
--- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Gly Arg Asp Phe Lys Arg Ala His His His His His His stop
```

FIG. 3D

Septavalent C.2

```
       M2--->
            9          18          27          36          45          54
5'  ATG AGT AAG AAC CCT GTC CCT GTC AAA AAA GAA GCA AAA TTA AGT GAA GCA GAA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Met Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu Ala Glu

           63          72          81          90          99         108
    TTA CAT GAC AAA ATT AAA AAC CTT GAA GAG GAA AAA GCA GAA TTA TTC GAG AAA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu Leu Phe Glu Lys

 (XhoI-SalI)M43--->
           117         126         135         144         153         162
    CTC GAC GAA GAA CAC CCT GAC GTT GTC GCT GCT AGA GAA AGC GTA CTA AAT AAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Leu Asp Glu Glu His Pro Asp Val Val Ala Ala Arg Glu Ser Val Leu Asn Asn

           171         180         189         198         207         216
    GTC CGT GTA CCG GGT ACA CTT TGG CTA CGT CAA AAA GAA GAA AAT GAC AAA CTT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Val Arg Val Pro Gly Thr Leu Trp Leu Arg Gln Lys Glu Glu Asn Asp Lys Leu

                                                                  NheI
           225         234         243         252         261         270
    AAA TTG GAA AAG AAA GGG CTT GAG ACT GAG TTA CAG GAA AAG GAA CAA GCT AGC
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Lys Leu Glu Lys Lys Gly Leu Glu Thr Glu Leu Gln Glu Lys Glu Gln Ala Ser

    M94--->
           279         288         297         306         315         324
    GAA GAA GCA TCA AAT AAT GGG CAA CTC ACA TTA CAG CAT AAA AAT AAT GCA TTG
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Glu Glu Ala Ser Asn Asn Gly Gln Leu Thr Leu Gln His Lys Asn Asn Ala Leu

           333         342         351         360         369         378
    ACT AGT GAG AAT GAG TCT CTT CGT CGT GAA AAA GAT CGT TAT TTG TAT GAA AAA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Thr Ser Glu Asn Glu Ser Leu Arg Arg Glu Lys Asp Arg Tyr Leu Tyr Glu Lys
```

FIG. 4A

```
                BamHI   M22--->
        387         396         405         414         423         432
GAA GAA TTA GAA GGA TCC GAG TCA TCA AAT AAT GCG GAG TCA TCA AAC ATT TCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Glu Leu Glu Gly Ser Glu Ser Ser Asn Asn Ala Glu Ser Ser Asn Ile Ser

        441         450         459         468         477         486
CAA GAA AGC AAA CTA ATA AAT ACA TTG ACT GAT GAA AAT GAG AAA CTC AGA GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Gln Glu Ser Lys Leu Ile Asn Thr Leu Thr Asp Glu Asn Glu Lys Leu Arg Glu

        495         504         513         522         531         540
GAG CTC CAA CAG TAT TAT GCA TTA AGT GAT GCT AAA GAA GAA GAA CCT CGT TAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Leu Gln Gln Tyr Tyr Ala Leu Ser Asp Ala Lys Glu Glu Glu Pro Arg Tyr

        PstI    M11--->
        549         558         567         576         585         594
AAA GCA CTG CAG ACT GAA GTT AAG GCT GCG GGG CAA AGC GCT CCT AAA GGT ACA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Lys Ala Leu Gln Thr Glu Val Lys Ala Ala Gly Gln Ser Ala Pro Lys Gly Thr

        603         612         621         630         639         648
AAC GTG AGC GCA GAC CTA TAT AAT TCG CTA TGG GAT GAA AAT AAA ACT CTT AGA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Val Ser Ala Asp Leu Tyr Asn Ser Leu Trp Asp Glu Asn Lys Thr Leu Arg

        657         666         675         684         693         702
GAA AAA CAA GAA GAG TAT ATA ACA AAA ATT CAA AAT GAA GAG ACA AAA AAT AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Lys Gln Glu Glu Tyr Ile Thr Lys Ile Gln Asn Glu Glu Thr Lys Asn Lys

KpnI    M59--->
        711         720         729         738         747         756
GGT ACC GAA CAA GCA AAA AAT AAT AAT GGG GAA CTC ACA TTA CAG CAA AAA TAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Gly Thr Glu Gln Ala Lys Asn Asn Asn Gly Glu Leu Thr Leu Gln Gln Lys Tyr

        765         774         783         792         801         810
GAT GCA TTG ACT AAT GAG AAT AAG TCT CTT CGT CGT GAG CGT GAT AAC TAT TTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asp Ala Leu Thr Asn Glu Asn Lys Ser Leu Arg Arg Glu Arg Asp Asn Tyr Leu
```

FIG. 4B

```
                          NcoI    M33--->
      819         828         837         846         855         864
AAT TAT TTA TAT GAA AAA CCA TGG GAA GAG CAT GAA AAA GTA ACA CAA GCC AGA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Tyr Leu Tyr Glu Lys Pro Trp Glu Glu His Glu Lys Val Thr Gln Ala Arg

      873         882         891         900         909         918
GAA GCG GTT ATC AGA GAG ATG CAA CAG AGG GGG ACA AAT TTT GGA CCT CTG TTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Ala Val Ile Arg Glu Met Gln Gln Arg Gly Thr Asn Phe Gly Pro Leu Leu

                                                                   PmlI
      927         936         945         954         963         972
GCA AGT ACA ATG CGA GAT AAT CAC AAT TTA AAA GAA ACG CTT GAC AAA ACT CAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Ser Thr Met Arg Asp Asn His Asn Leu Lys Glu Thr Leu Asp Lys Thr His

   M2--->
      981         990         999        1008        1017        1026
GTG AGT AAG AAC CCT GTC CCT GTC AAA AAA GAA GCA AAA TTA AGT GAA GCA GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Val Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu Ala Glu

     1035        1044        1053        1062        1071        1080
TTA CAT GAC AAA ATT AAA AAC CTT GAA GAG GAA AAA GCA GAA TTA TTC GAG AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu Leu Phe Glu Lys

     1089        1098        1107
CTC GAG CAC CAC CAC CAC CAC CAC TGA 3'
--- --- --- --- --- --- --- --- ---
Leu Glu His His His His His His stop
```

FIG. 4C

Septavalent D.3

```
       M89--->
            9          18          27          36          45          54
5'  ATG AGT GAC AAT ATT AAT CGT TCT GTC TCT GTC AAA GAT AAT GAA AAA GAA TTA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Met Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn Glu Lys Glu Leu

           63          72          81          90          99         108
    CAT AAC AAA ATT GCA GAC CTT GAA GAG GAA AGG GGT GAA CAT CTA GAC AAA ATA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His Leu Asp Lys Ile

                                                        BamHI   M101--->
          117         126         135         144         153         162
    GAT GAA CTA AAA GAA GAA CTA AAA GCA AAG GAA AAA AGT TCA GGA TCC GCT GAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser Gly Ser Ala Asp

          171         180         189         198         207         216
    CAC CCT AGC TAT ACC GCT GCT AAA GAT GAA GTA CTA AGT AAG TTC TCT GTA CCG
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    His Pro Ser Tyr Thr Ala Ala Lys Asp Glu Val Leu Ser Lys Phe Ser Val Pro

          225         234         243         252         261         270
    GGT CAT GTT TGG GCA CAT GAA AGA GAA AAA AAT GAC AAA CTT AGC TCG GAA AAT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Gly His Val Trp Ala His Glu Arg Glu Lys Asn Asp Lys Leu Ser Ser Glu Asn

                                                  NheI    M77--->
          279         288         297         306         315         324
    GAA GGG CTT AAG GCT GGT TTA CAG GAA AAG GAA CAA GCT AGC GAA GGG GTT TCT
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Glu Gly Leu Lys Ala Gly Leu Gln Glu Lys Glu Gln Ala Ser Glu Gly Val Ser

          333         342         351         360         369         378
    GTA GGT TCA GAT GCA TCA CTA CAT AAC CGC ATT ACA GAC CTT GAA GAG GAA AGA
    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
    Val Gly Ser Asp Ala Ser Leu His Asn Arg Ile Thr Asp Leu Glu Glu Glu Arg
```

FIG. 5A

```
        387         396         405         414         423         432
GAA AAA TTA TTA AAT AAA TTA GAT AAA GTT GAA GAA GAG CAT AAA AAA GAT CAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Lys Leu Leu Asn Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Asp His

        KpnI    M114--->
        441         450         459         468         477         486
GAA CAA GGT ACC AAC AGT AAG AAC CCT GCC CCT GCC CCT GCC TCT GCT GTC CCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Gln Gly Thr Asn Ser Lys Asn Pro Ala Pro Ala Pro Ala Ser Ala Val Pro

        495         504         513         522         531         540
GTC AAA AAA GAA GCA ACA AAA TTA AGT GAA GCA GAA TTA TAT AAC AAA ATT CAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Val Lys Lys Glu Ala Thr Lys Leu Ser Glu Ala Glu Leu Tyr Asn Lys Ile Gln

                                        BamHI   M75--->
        549         558         567         576         585         594
GAA CTT GAA GAG GGA AAA GCA GAA TTA TTC GGA TCC GAA GAA GAA CGT ACT TTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Leu Glu Glu Gly Lys Ala Glu Leu Phe Gly Ser Glu Glu Glu Arg Thr Phe

        603         612         621         630         639         648
ACT GAG TTA CCA TAT GAA GCA CGA TAC AAA GCA TGG AAA AGT GAA AAT GAT GAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Glu Leu Pro Tyr Glu Ala Arg Tyr Lys Ala Trp Lys Ser Glu Asn Asp Glu

        657         666         675         684         693         702
CTT CGG GAA AAT TAT CGT CGT ACC TTA GAT AAG TTT AAT ACT GAG CAA GGT AAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Arg Glu Asn Tyr Arg Arg Thr Leu Asp Lys Phe Asn Thr Glu Gln Gly Lys

                                HindIII M76--->
        711         720         729         738         747         756
ACT ACG CGC TTA GAA GAA CAA AAT AAG CTT GCG GAC GCG AAC TCG AAA AGC GTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Thr Arg Leu Glu Glu Gln Asn Lys Leu Ala Asp Ala Asn Ser Lys Ser Val

        765         774         783         792         801         810
TCT AAT AGT AAC GTG AGC ATA AAT CTA TAT AAT GAG CTA CAG GCT GAA CAT GAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ser Asn Ser Asn Val Ser Ile Asn Leu Tyr Asn Glu Leu Gln Ala Glu His Asp
```

FIG. 5B

```
        819         828         837         846         855         864
AAG CTA CAG ACT AAA CAT GAG GAG CTA TTG GCT GAA CAT GAT GCT CTT AAA GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Lys Leu Gln Thr Lys His Glu Glu Leu Leu Ala Glu His Asp Ala Leu Lys Glu

                        EcoRI   M92--->
        873         882         891         900         909         918
AAA CAA GAT AAA AAT CAA GAA TTC GAT GAC CGG AGC GTT TCT ACT AAT AGT GGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Lys Gln Asp Lys Asn Gln Glu Phe Asp Asp Arg Ser Val Ser Thr Asn Ser Gly

        927         936         945         954         963         972
AGC GTG AGC ACA CCA TAT AAT AAC CTA TTG AAT GAA TAT GAT GAC CTA TTG GCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ser Val Ser Thr Pro Tyr Asn Asn Leu Leu Asn Glu Tyr Asp Asp Leu Leu Ala

        981         990         999        1008        1017        1026
AAA CAT GGT GAG CTA TTG AGT GAA TAT GAT GCT CTT AAA GAA AAA CAA GAT AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Lys His Gly Glu Leu Leu Ser Glu Tyr Asp Ala Leu Lys Glu Lys Gln Asp Lys

            NheI    M89--->
       1035        1044        1053        1062        1071        1080
AAT CAA GAA GCT AGC AGT GAC AAT ATT AAT CGT TCT GTC TCT GTC AAA GAT AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Asn Gln Glu Ala Ser Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn

       1089        1098        1107        1116        1125        1134
GAA AAA GAA TTA CAT AAC AAA ATT GCA GAC CTT GAA GAG GAA AGG GGT GAA CAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Glu Lys Glu Leu His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His

       1143        1152        1161        1170        1179        1188
CTA GAC AAA ATA GAT GAA CTA AAA GAA GAA CTA AAA GCA AAG GAA AAA AGT TCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Asp Lys Ile Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser

       1197        1206
CAC CAC CAC CAC CAC CAC TGA 3'
--- --- --- --- --- --- ---
His His His His His His stop
```

FIG. 5C

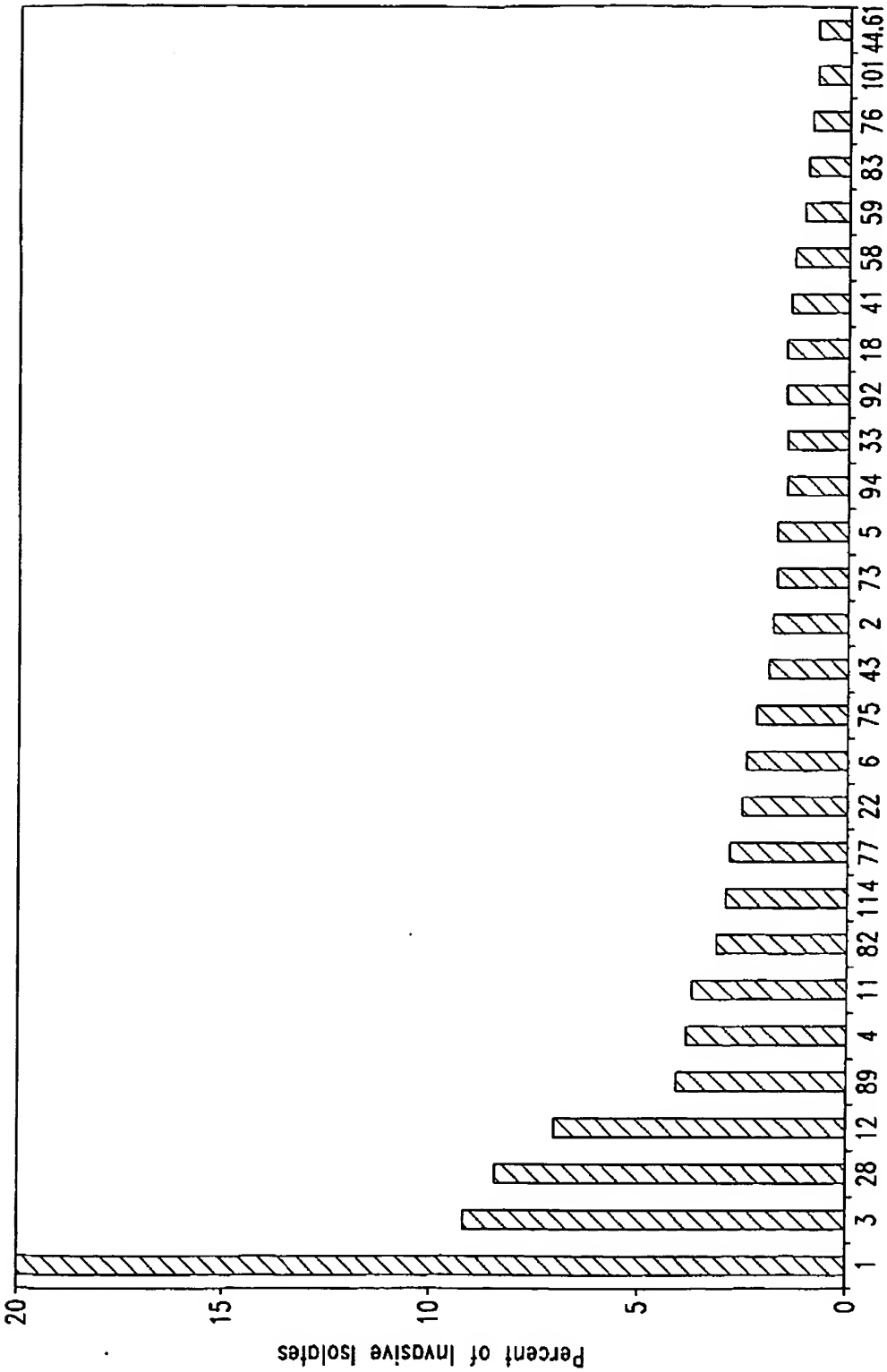

Percent of Invasive Isolates
0
5
10
15
20
M Type
1
3
28
12
89
4
11
82
114
77
22
6
75
43
2
73
5
94
33
92
18
41
58
59
83
76
101
44.61

FIG. 6

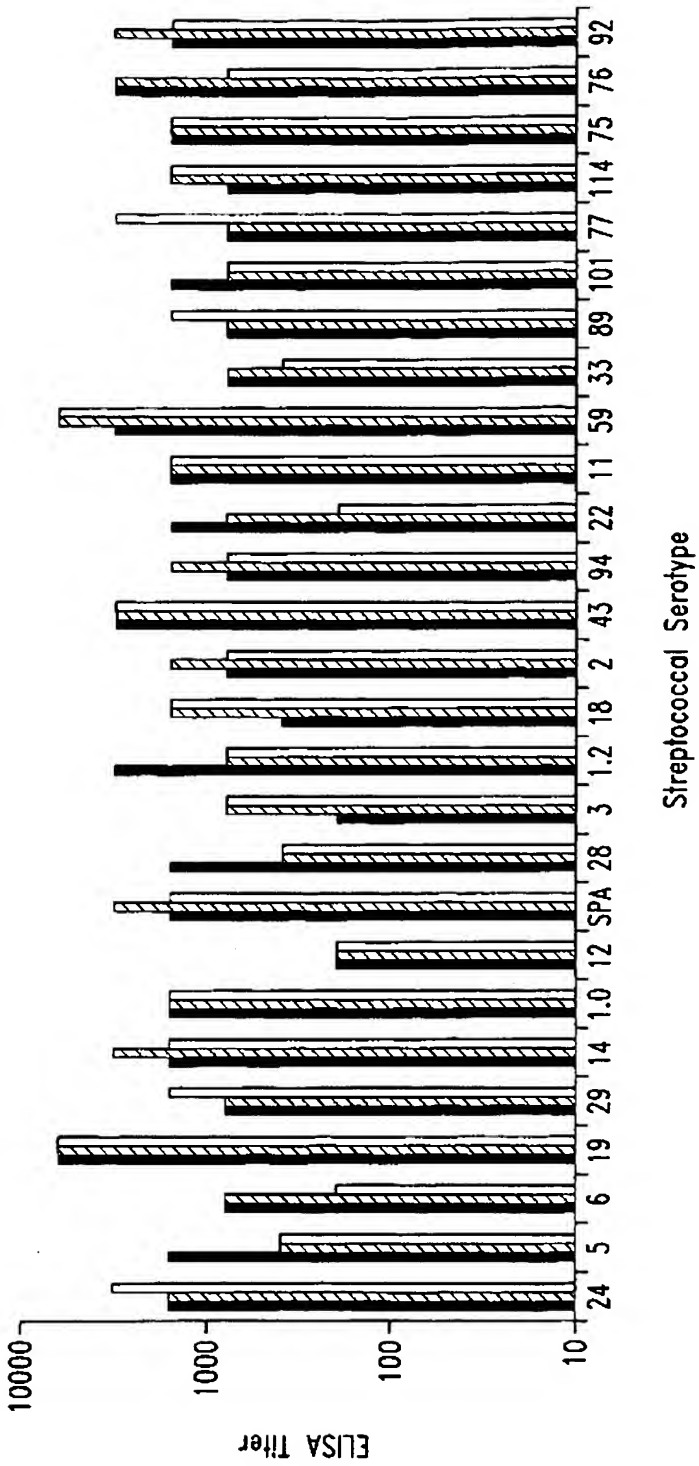

ELISA Titer
10000
1000
100
10
Streptococcal Serotype
24
5
6
19
29
14
1.0
12
SPA
28
3
1.2
18
2
43
94
22
11
59
33
89
101
77
114
75
76
92

FIG. 7

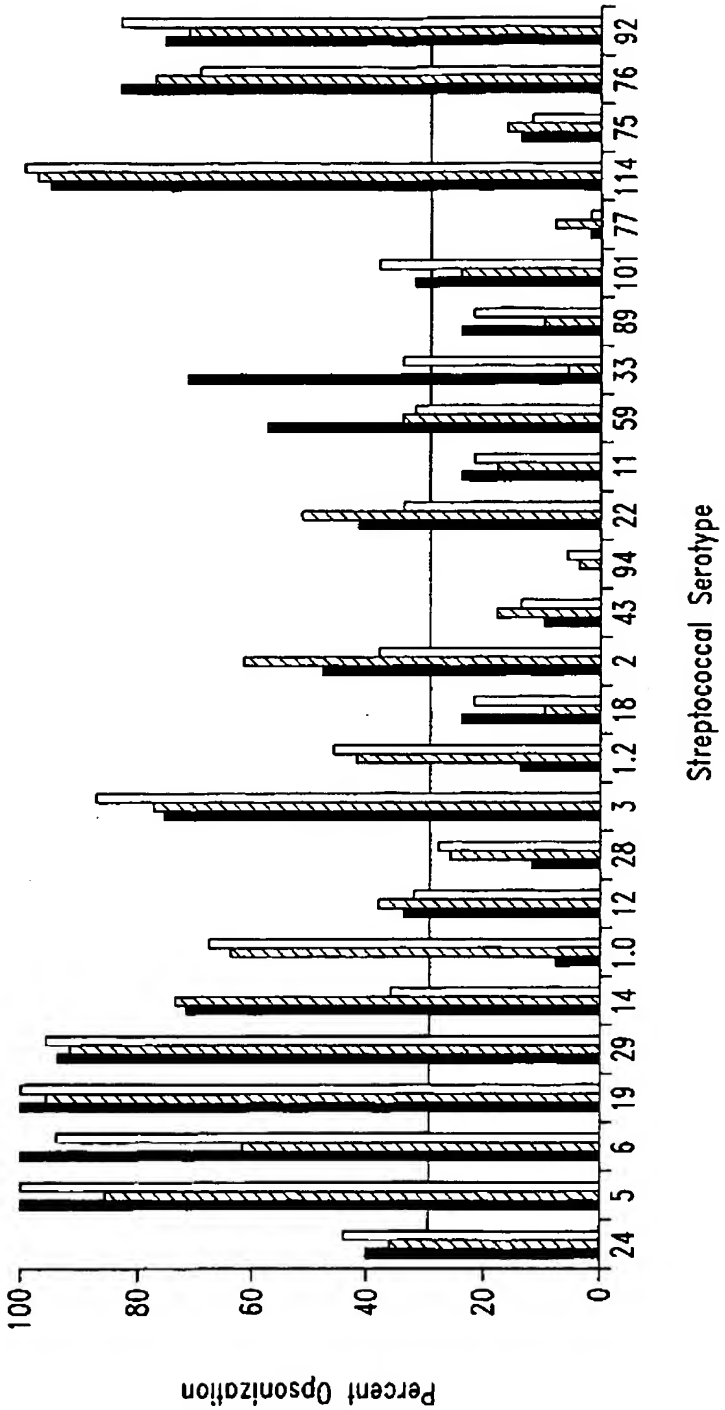

Percent Opsonization
0
20
40
60
80
100
24
5
6
19
29
14
1.0
12
28
3
1.2
18
2
43
94
22
11
59
33
89
101
77
114
75
76
92
Streptococcal Serotype

FIG. 8

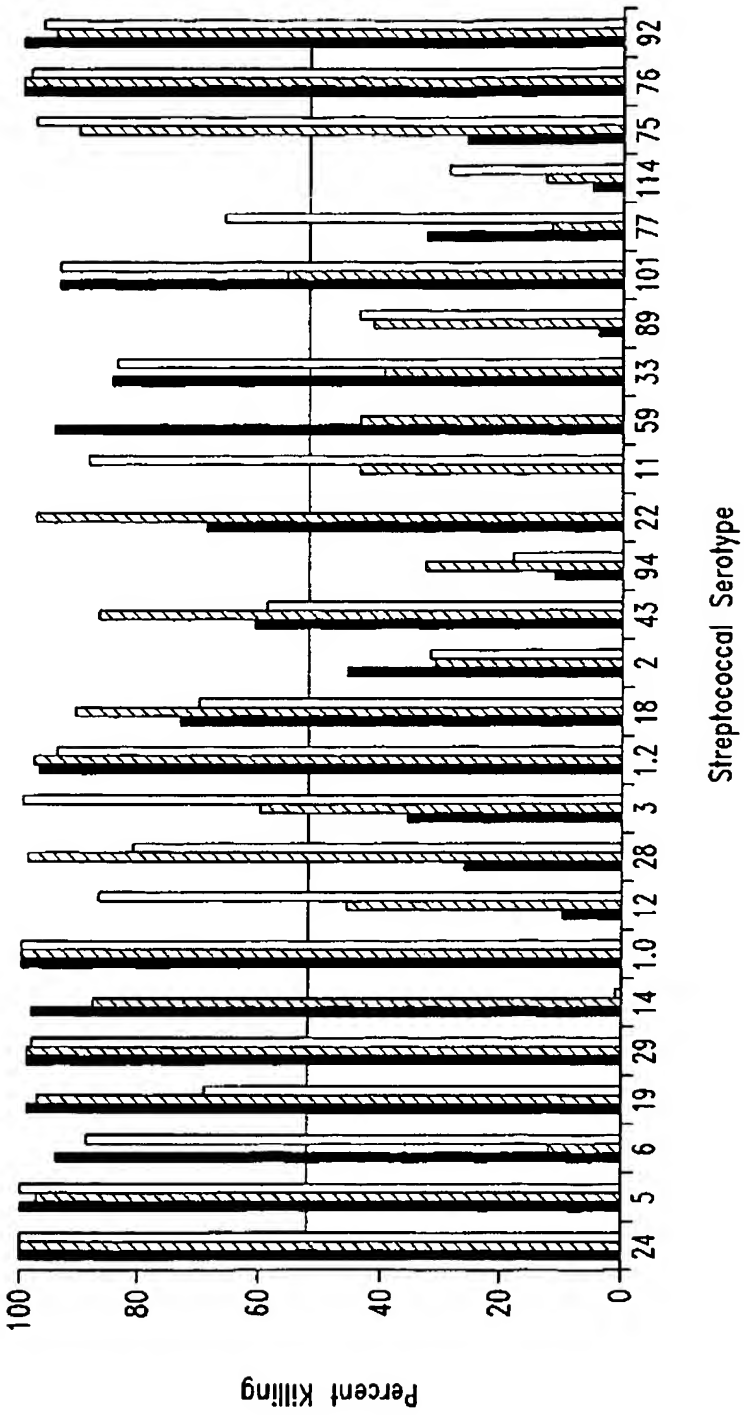

Percent Killing
100
80
60
40
20
0
24
5
6
19
29
14
1.0
12
28
3
1.2
18
2
43
94
22
11
59
33
89
101
77
114
75
76
92
Streptococcal Serotype

FIG. 9

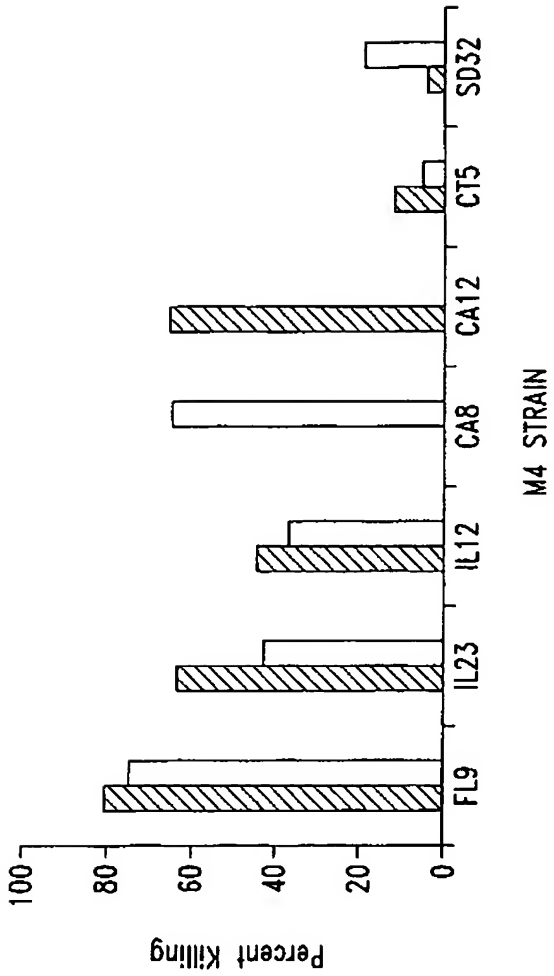

Percent Killing
100
80
60
40
20
0
FL9
IL23
IL12
CA8
CA12
CT5
SD32
M4 STRAIN

FIG. 10

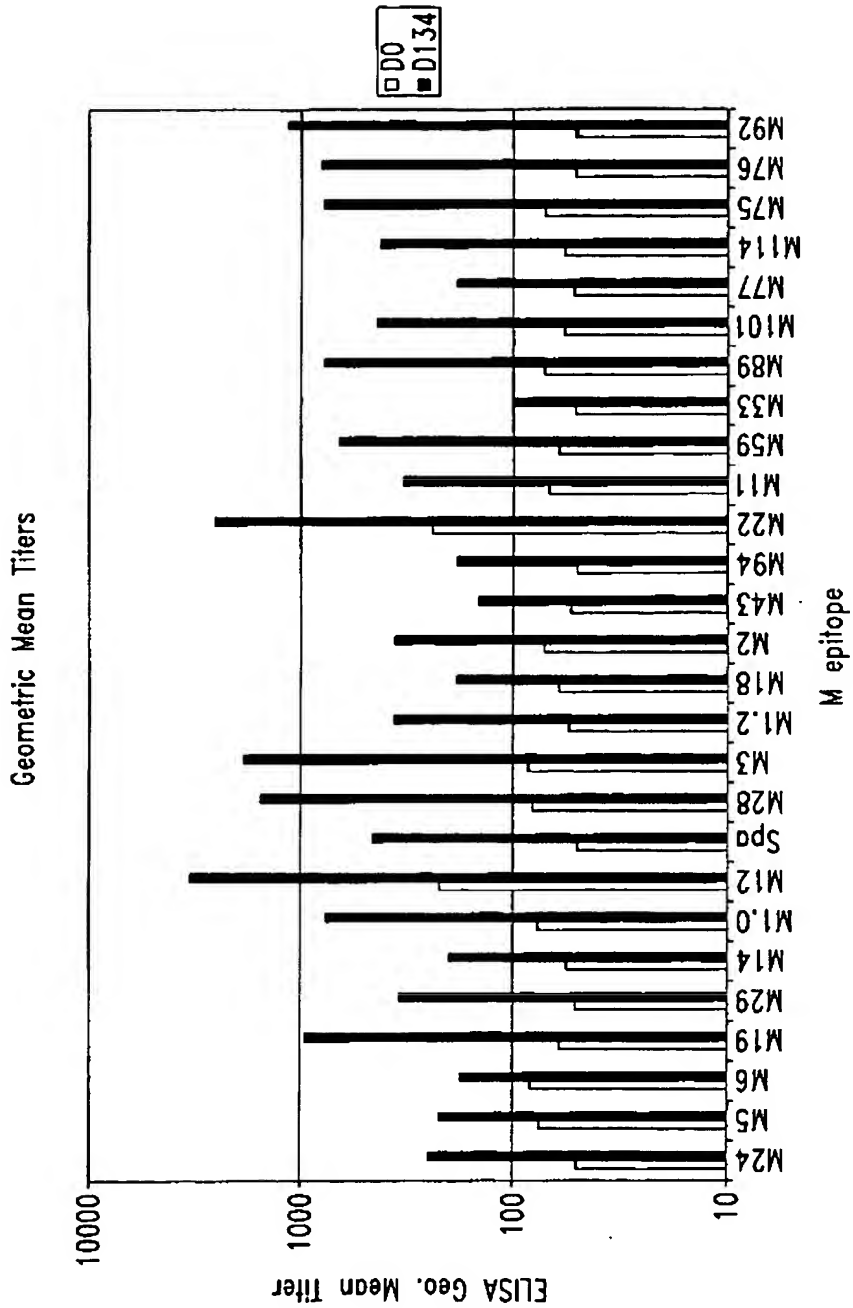

D0
D134
Geometric Mean Titers
ELISA Geo. Mean Titer
10000
1000
100
10
M24
M5
M6
M19
M29
M14
M1.0
M12
Spd
M28
M3
M1.2
M18
M2
M43
M94
M22
M11
M59
M33
M89
M101
M77
M114
M75
M76
M92
M epitope

FIG. 11

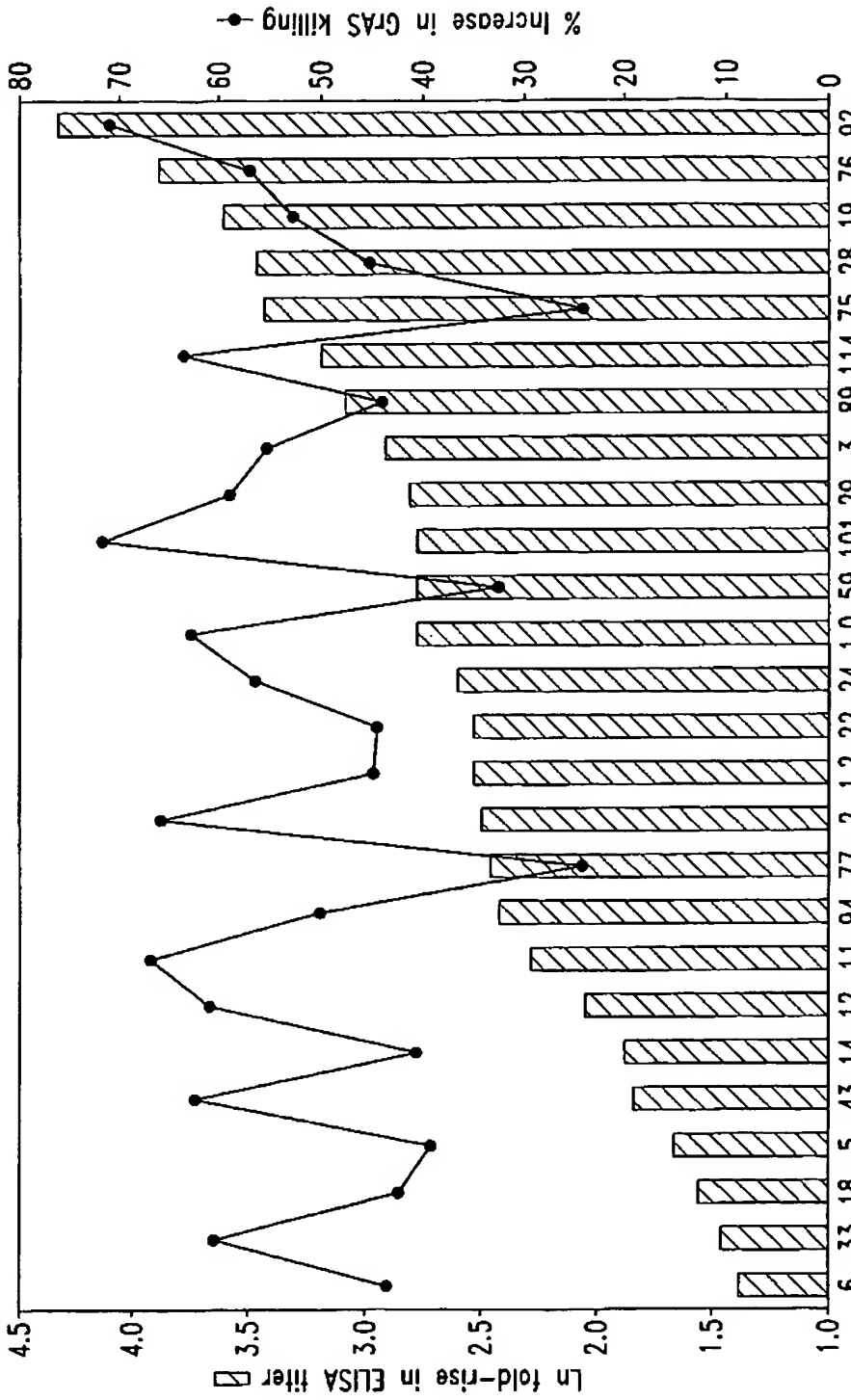

% Increase in GrAS killing
0
10
20
30
40
50
60
70
80
ln fold-rise in ELISA titer
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
6
33
18
5
43
14
12
11
94
77
2
1.2
22
24
1.0
59
101
29
3
89
114
75
28
19
76
92

FIG. 12

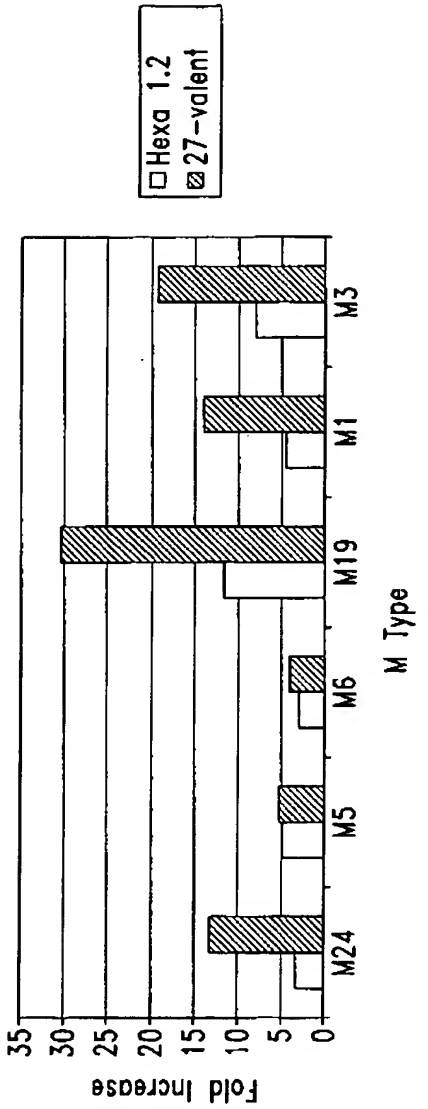

Hexa 1.2
27-valent
Fold Increase
0
5
10
15
20
25
30
35
M24
M5
M6
M19
M1
M3
M Type

FIG. 13

SEQUENCE LISTING

<110> ID Biomedical Corporation of Washington
University of Tennessee Research Corporation
Reddish, Mark A.
Hu, Mary C.
Walls, Michael A.
Dale, James B.

<120> MULTIVALENT STREPTOCOCCAL VACCINE COMPOSITIONS AND METHODS FOR USE

<130> 481112.413PC

<140> PCT
<141> 2002-10-28

<160> 19

<170> PatentIn Ver. 2.1

<210> 1
<211> 1005
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1002)

<220>
<223> Description of Artificial Sequence: Nucleic acid encoding Hybrid of Group A Streptococci M protein and M-like protein

<400> 1
atg gtc gcg act cgc tct cag aca gat act ctg gaa aaa gta caa gaa 48
Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu
1 5 10 15

cgt gct gac aag ttt gag ata gaa aac aat acg tta aaa ctt aag aat 96
Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn
20 25 30

agt gac tta agt ttt aat aat aaa gcg tta aaa gat cat aat gat gag 144
Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu
35 40 45

tta act gaa gag ttg agt aat gct aaa gag aaa cta cgt cac gtg gcc 192
Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg His Val Ala
50 55 60

gtg act cgc ggt aca ata aat gac ccg caa aga gca aaa gaa gct ctt 240
Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu Ala Leu
65 70 75 80

gac aag tat gag cta gaa aac cat gac tta aaa act aag gga tcc cgt 288

```
Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Gly Ser Arg
                 85                  90                  95

gtg ttt cct cgc ggg acg gta gaa aac ccg gac aaa gca cga gaa ctt   336
Val Phe Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu
            100                 105                 110

ctt aac aag tat gac gta gag aac tct atg tta caa gct aat aat gac   384
Leu Asn Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp
        115                 120                 125

aag tta cca tgg cgt gtg cgt tat act cgc cat acg cca gaa gat aag   432
Lys Leu Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys
    130                 135                 140

cta aaa aaa att att gac gat ctt gac gca aaa gaa cat gaa tta caa   480
Leu Lys Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln
145                 150                 155                 160

caa cag aat gag aag tta tct ctg cag aaa gtg tat att act cgt ggt   528
Gln Gln Asn Glu Lys Leu Ser Leu Gln Lys Val Tyr Ile Thr Arg Gly
                165                 170                 175

atg aca aaa gag gac gta gaa aaa att gct aac aac ctt gac ata gaa   576
Met Thr Lys Glu Asp Val Glu Lys Ile Ala Asn Asn Leu Asp Ile Glu
            180                 185                 190

aac cat ggg tta aaa caa cag aat gaa cag tta tct act gat aaa caa   624
Asn His Gly Leu Lys Gln Gln Asn Glu Gln Leu Ser Thr Asp Lys Gln
        195                 200                 205

ggt ctt gaa gaa cag aat ggt acc gat cgc gtt agt cgt tct atg tca   672
Gly Leu Glu Glu Gln Asn Gly Thr Asp Arg Val Ser Arg Ser Met Ser
    210                 215                 220

cgc gat gat cta tta aac agg gct cag gat ctt gaa gca aaa aac cac   720
Arg Asp Asp Leu Leu Asn Arg Ala Gln Asp Leu Glu Ala Lys Asn His
225                 230                 235                 240

ggg tta gaa cac cag aat act aag tta tct act gaa aat aaa acg ctt   768
Gly Leu Glu His Gln Asn Thr Lys Leu Ser Thr Glu Asn Lys Thr Leu
                245                 250                 255

caa gaa caa gca gaa gca cgc cag aaa gaa atc gat gtc gcg act cgc   816
Gln Glu Gln Ala Glu Ala Arg Gln Lys Glu Ile Asp Val Ala Thr Arg
            260                 265                 270

tct cag aca gat act ctg gaa aaa gta caa gaa cgt gct gac aag ttt   864
Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala Asp Lys Phe
        275                 280                 285

gag ata gaa aac aat acg tta aaa ctt aag aat agt gac tta agt ttt   912
Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser Phe
    290                 295                 300

aat aat aaa gcg tta aaa gat cat aat gat gag tta act gaa gag ttg   960
Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu
305                 310                 315                 320
```

agt aat gct aaa gag aaa cta cgt cac cac cac cac cac cac tga 1005
Ser Asn Ala Lys Glu Lys Leu Arg His His His His His His
325 330

<210> 2
<211> 334
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 2
Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu
1 5 10 15

Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn
20 25 30

Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu
35 40 45

Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg His Val Ala
50 55 60

Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu Ala Leu
65 70 75 80

Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Gly Ser Arg
85 90 95

Val Phe Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu
100 105 110

Leu Asn Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp
115 120 125

Lys Leu Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys
130 135 140

Leu Lys Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln
145 150 155 160

Gln Gln Asn Glu Lys Leu Ser Leu Gln Lys Val Tyr Ile Thr Arg Gly
165 170 175

Met Thr Lys Glu Asp Val Glu Lys Ile Ala Asn Asn Leu Asp Ile Glu
180 185 190

Asn His Gly Leu Lys Gln Gln Asn Glu Gln Leu Ser Thr Asp Lys Gln
195 200 205

Gly Leu Glu Glu Gln Asn Gly Thr Asp Arg Val Ser Arg Ser Met Ser
210 215 220

```
Arg Asp Asp Leu Leu Asn Arg Ala Gln Asp Leu Glu Ala Lys Asn His
225                 230                 235                 240

Gly Leu Glu His Gln Asn Thr Lys Leu Ser Thr Glu Asn Lys Thr Leu
                245                 250                 255

Gln Glu Gln Ala Glu Ala Arg Gln Lys Glu Ile Asp Val Ala Thr Arg
            260                 265                 270

Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala Asp Lys Phe
        275                 280                 285

Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser Phe
    290                 295                 300

Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu
305                 310                 315                 320

Ser Asn Ala Lys Glu Lys Leu Arg His His His His His His
                325                 330
```

<210> 3
<211> 1107
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1104)

<220>
<223> Description of Artificial Sequence: Hybrid of
Group A Streptococci M protein and M-like protein

<400> 3

```
atg agt aag aac cct gtc cct gtc aaa aaa gaa gca aaa tta agt gaa   48
Met Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu
  1               5                  10                  15

gca gaa tta cat gac aaa att aaa aac ctt gaa gag gaa aaa gca gaa   96
Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu
             20                  25                  30

tta ttc gag aaa ctc gac gaa gaa cac cct gac gtt gtc gct gct aga   144
Leu Phe Glu Lys Leu Asp Glu Glu His Pro Asp Val Val Ala Ala Arg
         35                  40                  45

gaa agc gta cta aat aat gtc cgt gta ccg ggt aca ctt tgg cta cgt   192
Glu Ser Val Leu Asn Asn Val Arg Val Pro Gly Thr Leu Trp Leu Arg
     50                  55                  60

caa aaa gaa gaa aat gac aaa ctt aaa ttg gaa aag aaa ggg ctt gag   240
Gln Lys Glu Glu Asn Asp Lys Leu Lys Leu Glu Lys Lys Gly Leu Glu
 65                  70                  75                  80

act gag tta cag gaa aag gaa caa gct agc gaa gaa gca tca aat aat   288
```

```
Thr Glu Leu Gln Glu Lys Glu Gln Ala Ser Glu Glu Ala Ser Asn Asn
                 85                  90                  95

ggg caa ctc aca tta cag cat aaa aat aat gca ttg act agt gag aat   336
Gly Gln Leu Thr Leu Gln His Lys Asn Asn Ala Leu Thr Ser Glu Asn
            100                 105                 110

gag tct ctt cgt cgt gaa aaa gat cgt tat ttg tat gaa aaa gaa gaa   384
Glu Ser Leu Arg Arg Glu Lys Asp Arg Tyr Leu Tyr Glu Lys Glu Glu
        115                 120                 125

tta gaa gga tcc gag tca tca aat aat gcg gag tca tca aac att tct   432
Leu Glu Gly Ser Glu Ser Ser Asn Asn Ala Glu Ser Ser Asn Ile Ser
    130                 135                 140

caa gaa agc aaa cta ata aat aca ttg act gat gaa aat gag aaa ctc   480
Gln Glu Ser Lys Leu Ile Asn Thr Leu Thr Asp Glu Asn Glu Lys Leu
145                 150                 155                 160

aga gaa gag ctc caa cag tat tat gca tta agt gat gct aaa gaa gaa   528
Arg Glu Glu Leu Gln Gln Tyr Tyr Ala Leu Ser Asp Ala Lys Glu Glu
                165                 170                 175

gaa cct cgt tat aaa gca ctg cag act gaa gtt aag gct gcg ggg caa   576
Glu Pro Arg Tyr Lys Ala Leu Gln Thr Glu Val Lys Ala Ala Gly Gln
            180                 185                 190

agc gct cct aaa ggt aca aac gtg agc gca gac cta tat aat tcg cta   624
Ser Ala Pro Lys Gly Thr Asn Val Ser Ala Asp Leu Tyr Asn Ser Leu
        195                 200                 205

tgg gat gaa aat aaa act ctt aga gaa aaa caa gaa gag tat ata aca   672
Trp Asp Glu Asn Lys Thr Leu Arg Glu Lys Gln Glu Glu Tyr Ile Thr
    210                 215                 220

aaa att caa aat gaa gag aca aaa aat aaa ggt acc gaa caa gca aaa   720
Lys Ile Gln Asn Glu Glu Thr Lys Asn Lys Gly Thr Glu Gln Ala Lys
225                 230                 235                 240

aat aat aat ggg gaa ctc aca tta cag caa aaa tac gat gca ttg act   768
Asn Asn Asn Gly Glu Leu Thr Leu Gln Gln Lys Tyr Asp Ala Leu Thr
                245                 250                 255

aat gag aat aag tct ctt cgt cgt gag cgt gat aac tat tta aat tat   816
Asn Glu Asn Lys Ser Leu Arg Arg Glu Arg Asp Asn Tyr Leu Asn Tyr
            260                 265                 270

tta tat gaa aaa cca tgg gaa gag cat gaa aaa gta aca caa gcc aga   864
Leu Tyr Glu Lys Pro Trp Glu Glu His Glu Lys Val Thr Gln Ala Arg
        275                 280                 285

gaa gcg gtt atc aga gag atg caa cag agg ggg aca aat ttt gga cct   912
Glu Ala Val Ile Arg Glu Met Gln Gln Arg Gly Thr Asn Phe Gly Pro
    290                 295                 300

ctg tta gca agt aca atg cga gat aat cac aat tta aaa gaa acg ctt   960
Leu Leu Ala Ser Thr Met Arg Asp Asn His Asn Leu Lys Glu Thr Leu
305                 310                 315                 320
```

```
gac aaa act cac gtg agt aag aac cct gtc cct gtc aaa aaa gaa gca   1008
Asp Lys Thr His Val Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala
                325                 330                 335

aaa tta agt gaa gca gaa tta cat gac aaa att aaa aac ctt gaa gag   1056
Lys Leu Ser Glu Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu
            340                 345                 350

gaa aaa gca gaa tta ttc gag aaa ctc gag cac cac cac cac cac cac   1104
Glu Lys Ala Glu Leu Phe Glu Lys Leu Glu His His His His His His
        355                 360                 365

tga                                                              1107
```

<210> 4
<211> 368
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 4

```
Met Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu
  1               5                  10                  15

Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu
             20                  25                  30

Leu Phe Glu Lys Leu Asp Glu Glu His Pro Asp Val Val Ala Ala Arg
         35                  40                  45

Glu Ser Val Leu Asn Asn Val Arg Val Pro Gly Thr Leu Trp Leu Arg
     50                  55                  60

Gln Lys Glu Glu Asn Asp Lys Leu Lys Leu Glu Lys Lys Gly Leu Glu
 65                  70                  75                  80

Thr Glu Leu Gln Glu Lys Glu Gln Ala Ser Glu Glu Ala Ser Asn Asn
                 85                  90                  95

Gly Gln Leu Thr Leu Gln His Lys Asn Asn Ala Leu Thr Ser Glu Asn
            100                 105                 110

Glu Ser Leu Arg Arg Glu Lys Asp Arg Tyr Leu Tyr Glu Lys Glu Glu
        115                 120                 125

Leu Glu Gly Ser Glu Ser Ser Asn Asn Ala Glu Ser Ser Asn Ile Ser
    130                 135                 140

Gln Glu Ser Lys Leu Ile Asn Thr Leu Thr Asp Glu Asn Glu Lys Leu
145                 150                 155                 160

Arg Glu Glu Leu Gln Gln Tyr Tyr Ala Leu Ser Asp Ala Lys Glu Glu
                165                 170                 175
```

```
Glu Pro Arg Tyr Lys Ala Leu Gln Thr Glu Val Lys Ala Ala Gly Gln
            180                 185                 190

Ser Ala Pro Lys Gly Thr Asn Val Ser Ala Asp Leu Tyr Asn Ser Leu
        195                 200                 205

Trp Asp Glu Asn Lys Thr Leu Arg Glu Lys Gln Glu Glu Tyr Ile Thr
    210                 215                 220

Lys Ile Gln Asn Glu Glu Thr Lys Asn Lys Gly Thr Glu Gln Ala Lys
225                 230                 235                 240

Asn Asn Asn Gly Glu Leu Thr Leu Gln Gln Lys Tyr Asp Ala Leu Thr
                245                 250                 255

Asn Glu Asn Lys Ser Leu Arg Arg Glu Arg Asp Asn Tyr Leu Asn Tyr
            260                 265                 270

Leu Tyr Glu Lys Pro Trp Glu Glu His Glu Lys Val Thr Gln Ala Arg
        275                 280                 285

Glu Ala Val Ile Arg Glu Met Gln Gln Arg Gly Thr Asn Phe Gly Pro
    290                 295                 300

Leu Leu Ala Ser Thr Met Arg Asp Asn His Asn Leu Lys Glu Thr Leu
305                 310                 315                 320

Asp Lys Thr His Val Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala
                325                 330                 335

Lys Leu Ser Glu Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu
            340                 345                 350

Glu Lys Ala Glu Leu Phe Glu Lys Leu Glu His His His His His His
        355                 360                 365
```

<210> 5
<211> 1209
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1206)

<220>
<223> Description of Artificial Sequence: Nucleic acid encoding Hybrid of Group A Streptococci M protein and M-like protein

<400> 5

```
atg agt gac aat att aat cgt tct gtc tct gtc aaa gat aat gaa aaa   48
Met Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn Glu Lys
  1               5                  10                  15
```

```
gaa tta cat aac aaa att gca gac ctt gaa gag gaa agg ggt gaa cat   96
Glu Leu His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His
            20                  25                  30

cta gac aaa ata gat gaa cta aaa gaa gaa cta aaa gca aag gaa aaa  144
Leu Asp Lys Ile Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys
        35                  40                  45

agt tca gga tcc gct gat cac cct agc tat acc gct gct aaa gat gaa  192
Ser Ser Gly Ser Ala Asp His Pro Ser Tyr Thr Ala Ala Lys Asp Glu
    50                  55                  60

gta cta agt aag ttc tct gta ccg ggt cat gtt tgg gca cat gaa aga  240
Val Leu Ser Lys Phe Ser Val Pro Gly His Val Trp Ala His Glu Arg
65                  70                  75                  80

gaa aaa aat gac aaa ctt agc tcg gaa aat gaa ggg ctt aag gct ggt  288
Glu Lys Asn Asp Lys Leu Ser Ser Glu Asn Glu Gly Leu Lys Ala Gly
                85                  90                  95

tta cag gaa aag gaa caa gct agc gaa ggg gtt tct gta ggt tca gat  336
Leu Gln Glu Lys Glu Gln Ala Ser Glu Gly Val Ser Val Gly Ser Asp
            100                 105                 110

gca tca cta cat aac cgc att aca gac ctt gaa gag gaa aga gaa aaa  384
Ala Ser Leu His Asn Arg Ile Thr Asp Leu Glu Glu Glu Arg Glu Lys
        115                 120                 125

tta tta aat aaa tta gat aaa gtt gaa gaa gag cat aaa aaa gat cat  432
Leu Leu Asn Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Asp His
    130                 135                 140

gaa caa ggt acc aac agt aag aac cct gcc cct gcc cct gcc tct gct  480
Glu Gln Gly Thr Asn Ser Lys Asn Pro Ala Pro Ala Pro Ala Ser Ala
145                 150                 155                 160

gtc cct gtc aaa aaa gaa gca aca aaa tta agt gaa gca gaa tta tat  528
Val Pro Val Lys Lys Glu Ala Thr Lys Leu Ser Glu Ala Glu Leu Tyr
                165                 170                 175

aac aaa att caa gaa ctt gaa gag gga aaa gca gaa tta ttc gga tcc  576
Asn Lys Ile Gln Glu Leu Glu Glu Gly Lys Ala Glu Leu Phe Gly Ser
            180                 185                 190

gaa gaa gaa cgt act ttt act gag tta cca tat gaa gca cga tac aaa  624
Glu Glu Glu Arg Thr Phe Thr Glu Leu Pro Tyr Glu Ala Arg Tyr Lys
        195                 200                 205

gca tgg aaa agt gaa aat gat gag ctt cgg gaa aat tat cgt cgt acc  672
Ala Trp Lys Ser Glu Asn Asp Glu Leu Arg Glu Asn Tyr Arg Arg Thr
    210                 215                 220

tta gat aag ttt aat act gag caa ggt aag act acg cgc tta gaa gaa  720
Leu Asp Lys Phe Asn Thr Glu Gln Gly Lys Thr Thr Arg Leu Glu Glu
225                 230                 235                 240

caa aat aag ctt gcg gac gcg aac tcg aaa agc gtt tct aat agt aac  768
Gln Asn Lys Leu Ala Asp Ala Asn Ser Lys Ser Val Ser Asn Ser Asn
```

```
                245                 250                 255

gtg agc ata aat cta tat aat gag cta cag gct gaa cat gat aag cta   816
Val Ser Ile Asn Leu Tyr Asn Glu Leu Gln Ala Glu His Asp Lys Leu
            260                 265                 270

cag act aaa cat gag gag cta ttg gct gaa cat gat gct ctt aaa gaa   864
Gln Thr Lys His Glu Glu Leu Leu Ala Glu His Asp Ala Leu Lys Glu
        275                 280                 285

aaa caa gat aaa aat caa gaa ttc gat gac cgg agc gtt tct act aat   912
Lys Gln Asp Lys Asn Gln Glu Phe Asp Asp Arg Ser Val Ser Thr Asn
    290                 295                 300

agt ggt agc gtg agc aca cca tat aat aac cta ttg aat gaa tat gat   960
Ser Gly Ser Val Ser Thr Pro Tyr Asn Asn Leu Leu Asn Glu Tyr Asp
305                 310                 315                 320

gac cta ttg gct aaa cat ggt gag cta ttg agt gaa tat gat gct ctt   1008
Asp Leu Leu Ala Lys His Gly Glu Leu Leu Ser Glu Tyr Asp Ala Leu
                325                 330                 335

aaa gaa aaa caa gat aaa aat caa gaa gct agc agt gac aat att aat   1056
Lys Glu Lys Gln Asp Lys Asn Gln Glu Ala Ser Ser Asp Asn Ile Asn
            340                 345                 350

cgt tct gtc tct gtc aaa gat aat gaa aaa gaa tta cat aac aaa att   1104
Arg Ser Val Ser Val Lys Asp Asn Glu Lys Glu Leu His Asn Lys Ile
        355                 360                 365

gca gac ctt gaa gag gaa agg ggt gaa cat cta gac aaa ata gat gaa   1152
Ala Asp Leu Glu Glu Glu Arg Gly Glu His Leu Asp Lys Ile Asp Glu
    370                 375                 380

cta aaa gaa gaa cta aaa gca aag gaa aaa agt tca cac cac cac cac   1200
Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser His His His His
385                 390                 395                 400

cac cac tga                                                      1209
His His
```

<210> 6
<211> 402
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 6

```
Met Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn Glu Lys
  1               5                  10                  15

Glu Leu His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His
             20                  25                  30
```

```
Leu Asp Lys Ile Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys
        35                  40                  45

Ser Ser Gly Ser Ala Asp His Pro Ser Tyr Thr Ala Ala Lys Asp Glu
    50                  55                  60

Val Leu Ser Lys Phe Ser Val Pro Gly His Val Trp Ala His Glu Arg
65                  70                  75                  80

Glu Lys Asn Asp Lys Leu Ser Ser Glu Asn Glu Gly Leu Lys Ala Gly
                85                  90                  95

Leu Gln Glu Lys Glu Gln Ala Ser Glu Gly Val Ser Val Gly Ser Asp
            100                 105                 110

Ala Ser Leu His Asn Arg Ile Thr Asp Leu Glu Glu Glu Arg Glu Lys
        115                 120                 125

Leu Leu Asn Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Asp His
    130                 135                 140

Glu Gln Gly Thr Asn Ser Lys Asn Pro Ala Pro Ala Pro Ala Ser Ala
145                 150                 155                 160

Val Pro Val Lys Lys Glu Ala Thr Lys Leu Ser Glu Ala Glu Leu Tyr
                165                 170                 175

Asn Lys Ile Gln Glu Leu Glu Glu Gly Lys Ala Glu Leu Phe Gly Ser
            180                 185                 190

Glu Glu Glu Arg Thr Phe Thr Glu Leu Pro Tyr Glu Ala Arg Tyr Lys
        195                 200                 205

Ala Trp Lys Ser Glu Asn Asp Glu Leu Arg Glu Asn Tyr Arg Arg Thr
    210                 215                 220

Leu Asp Lys Phe Asn Thr Glu Gln Gly Lys Thr Thr Arg Leu Glu Glu
225                 230                 235                 240

Gln Asn Lys Leu Ala Asp Ala Asn Ser Lys Ser Val Ser Asn Ser Asn
                245                 250                 255

Val Ser Ile Asn Leu Tyr Asn Glu Leu Gln Ala Glu His Asp Lys Leu
            260                 265                 270

Gln Thr Lys His Glu Glu Leu Leu Ala Glu His Asp Ala Leu Lys Glu
        275                 280                 285

Lys Gln Asp Lys Asn Gln Glu Phe Asp Asp Arg Ser Val Ser Thr Asn
    290                 295                 300

Ser Gly Ser Val Ser Thr Pro Tyr Asn Asn Leu Leu Asn Glu Tyr Asp
305                 310                 315                 320

Asp Leu Leu Ala Lys His Gly Glu Leu Leu Ser Glu Tyr Asp Ala Leu
                325                 330                 335

Lys Glu Lys Gln Asp Lys Asn Gln Glu Ala Ser Ser Asp Asn Ile Asn
```

```
            340                 345                 350

Arg Ser Val Ser Val Lys Asp Asn Glu Lys Glu Leu His Asn Lys Ile
        355                 360                 365

Ala Asp Leu Glu Glu Glu Arg Gly Glu His Leu Asp Lys Ile Asp Glu
    370                 375                 380

Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser His His His His
385                 390                 395                 400

His His
```

<210> 7
<211> 1287
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1284)

<220>
<223> Description of Artificial Sequence: Hybrid of
Group A Streptococci M protein and M-like protein

<400> 7

```
atg aac ggt gat ggt aat cct agg gaa gtt ata gaa gat ctt gca gca   48
Met Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala
  1               5                  10                  15

aac aat ccc gca ata caa aat ata cgt tta cgt cac gaa aac aag gac   96
Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp
             20                  25                  30

tta aaa gcg aga tta gag aat gca atg gaa gtt gca gga aga gat ttt   144
Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe
         35                  40                  45

aag aga gct ctc gac gat cat agt gat tta gtc gca gaa aaa caa cgt   192
Lys Arg Ala Leu Asp Asp His Ser Asp Leu Val Ala Glu Lys Gln Arg
     50                  55                  60

tta gaa gat tta gga caa aaa ttt gaa aga ctg aaa cag cgt tca gaa   240
Leu Glu Asp Leu Gly Gln Lys Phe Glu Arg Leu Lys Gln Arg Ser Glu
 65                  70                  75                  80

ctc tac ctt cag caa tac tat gat aat aaa tca aat gga tat aaa ggt   288
Leu Tyr Leu Gln Gln Tyr Tyr Asp Asn Lys Ser Asn Gly Tyr Lys Gly
                 85                  90                  95

gac tgg tat gta caa cag tta gga tcc gat tca gta agt gga tta gag   336
Asp Trp Tyr Val Gln Gln Leu Gly Ser Asp Ser Val Ser Gly Leu Glu
            100                 105                 110
```

```
gtg gca gac ccc tct gat agt aag aaa ctt att gaa tta ggt ttg gct   384
Val Ala Asp Pro Ser Asp Ser Lys Lys Leu Ile Glu Leu Gly Leu Ala
        115                 120                 125

aaa tac ctt aat gat aaa tta ccc ttt aaa act aaa gaa gat tca gag   432
Lys Tyr Leu Asn Asp Lys Leu Pro Phe Lys Thr Lys Glu Asp Ser Glu
    130                 135                 140

att tta tca gag tta cgt gat gta tta aaa aat ctg cag gag tct cca   480
Ile Leu Ser Glu Leu Arg Asp Val Leu Lys Asn Leu Gln Glu Ser Pro
145                 150                 155                 160

aaa agt act gag act tct gct aat gga gct gat aaa tta gct gat gca   528
Lys Ser Thr Glu Thr Ser Ala Asn Gly Ala Asp Lys Leu Ala Asp Ala
                165                 170                 175

tac aac aca ttg ctt act gaa cat gag aaa ctc aga gat gag tat tat   576
Tyr Asn Thr Leu Leu Thr Glu His Glu Lys Leu Arg Asp Glu Tyr Tyr
            180                 185                 190

aca tta att gat gct aaa gaa gaa gaa cct cgc tat aaa gca ttg ggt   624
Thr Leu Ile Asp Ala Lys Glu Glu Glu Pro Arg Tyr Lys Ala Leu Gly
        195                 200                 205

acc ttg tta gat cag gtt aca caa tta tat act aaa cat aat agt aat   672
Thr Leu Leu Asp Gln Val Thr Gln Leu Tyr Thr Lys His Asn Ser Asn
    210                 215                 220

tac caa caa tat aat gca caa gct ggc aga ctt gac ctg aga caa aag   720
Tyr Gln Gln Tyr Asn Ala Gln Ala Gly Arg Leu Asp Leu Arg Gln Lys
225                 230                 235                 240

gct gaa tat cta aaa ggc ctt aat gat tgg gct gag cgc ctg tta caa   768
Ala Glu Tyr Leu Lys Gly Leu Asn Asp Trp Ala Glu Arg Leu Leu Gln
                245                 250                 255

gag tta aat ggt acc aac aat gat ggt cgt tct cgt gac gtt acg gaa   816
Glu Leu Asn Gly Thr Asn Asn Asp Gly Arg Ser Arg Asp Val Thr Glu
            260                 265                 270

gag att gca gca aac aat acc aca gta caa aat ata cgt tta cgt aac   864
Glu Ile Ala Ala Asn Asn Thr Thr Val Gln Asn Ile Arg Leu Arg Asn
        275                 280                 285

gaa aac aag aac tta aaa gcg aaa aac gag gac tta gaa gcg aga tta   912
Glu Asn Lys Asn Leu Lys Ala Lys Asn Glu Asp Leu Glu Ala Arg Leu
    290                 295                 300

gag aat gca atg aat gtt gca gga cgc gat ttt aag cgt gct gaa ttc   960
Glu Asn Ala Met Asn Val Ala Gly Arg Asp Phe Lys Arg Ala Glu Phe
305                 310                 315                 320

gca cct ctt act cgt gct aca gca gac aat aaa gac gaa tta ata aaa  1008
Ala Pro Leu Thr Arg Ala Thr Ala Asp Asn Lys Asp Glu Leu Ile Lys
                325                 330                 335

aga gct aac ggt tat gag ata cag aac cat cag tta aca gtt gag aat  1056
Arg Ala Asn Gly Tyr Glu Ile Gln Asn His Gln Leu Thr Val Glu Asn
```

```
            340                 345                 350

aaa aaa tta aaa act gat aag gaa cag tta aca aaa gag aat gat gat   1104
Lys Lys Leu Lys Thr Asp Lys Glu Gln Leu Thr Lys Glu Asn Asp Asp
        355                 360                 365

tta aaa cac gtg aac ggt gat ggt aat cct cgt gaa gtt ata gaa gat   1152
Leu Lys His Val Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp
    370                 375                 380

ctt gca gca aac aat ccc gca ata caa aat ata cgt tta cgt cac gaa   1200
Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu
385                 390                 395                 400

aac aag gac tta aaa gcg aga tta gag aat gca atg gaa gtt gca gga   1248
Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly
                405                 410                 415

cgt gat ttt aag cgt gct cac cac cac cac cac cac taa              1287
Arg Asp Phe Lys Arg Ala His His His His His His
            420                 425


<210> 8
<211> 428
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M
      protein and M-like protein

<400> 8
Met Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala
  1               5                  10                  15

Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp
             20                  25                  30

Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe
         35                  40                  45

Lys Arg Ala Leu Asp Asp His Ser Asp Leu Val Ala Glu Lys Gln Arg
     50                  55                  60

Leu Glu Asp Leu Gly Gln Lys Phe Glu Arg Leu Lys Gln Arg Ser Glu
 65                  70                  75                  80

Leu Tyr Leu Gln Gln Tyr Tyr Asp Asn Lys Ser Asn Gly Tyr Lys Gly
                 85                  90                  95

Asp Trp Tyr Val Gln Gln Leu Gly Ser Asp Ser Val Ser Gly Leu Glu
            100                 105                 110

Val Ala Asp Pro Ser Asp Ser Lys Lys Leu Ile Glu Leu Gly Leu Ala
        115                 120                 125

Lys Tyr Leu Asn Asp Lys Leu Pro Phe Lys Thr Lys Glu Asp Ser Glu
```

```
    130                 135                 140

Ile Leu Ser Glu Leu Arg Asp Val Leu Lys Asn Leu Gln Glu Ser Pro
145                 150                 155                 160

Lys Ser Thr Glu Thr Ser Ala Asn Gly Ala Asp Lys Leu Ala Asp Ala
                165                 170                 175

Tyr Asn Thr Leu Leu Thr Glu His Glu Lys Leu Arg Asp Glu Tyr Tyr
            180                 185                 190

Thr Leu Ile Asp Ala Lys Glu Glu Glu Pro Arg Tyr Lys Ala Leu Gly
        195                 200                 205

Thr Leu Leu Asp Gln Val Thr Gln Leu Tyr Thr Lys His Asn Ser Asn
    210                 215                 220

Tyr Gln Gln Tyr Asn Ala Gln Ala Gly Arg Leu Asp Leu Arg Gln Lys
225                 230                 235                 240

Ala Glu Tyr Leu Lys Gly Leu Asn Asp Trp Ala Glu Arg Leu Leu Gln
                245                 250                 255

Glu Leu Asn Gly Thr Asn Asn Asp Gly Arg Ser Arg Asp Val Thr Glu
            260                 265                 270

Glu Ile Ala Ala Asn Asn Thr Thr Val Gln Asn Ile Arg Leu Arg Asn
        275                 280                 285

Glu Asn Lys Asn Leu Lys Ala Lys Asn Glu Asp Leu Glu Ala Arg Leu
    290                 295                 300

Glu Asn Ala Met Asn Val Ala Gly Arg Asp Phe Lys Arg Ala Glu Phe
305                 310                 315                 320

Ala Pro Leu Thr Arg Ala Thr Ala Asp Asn Lys Asp Glu Leu Ile Lys
                325                 330                 335

Arg Ala Asn Gly Tyr Glu Ile Gln Asn His Gln Leu Thr Val Glu Asn
            340                 345                 350

Lys Lys Leu Lys Thr Asp Lys Glu Gln Leu Thr Lys Glu Asn Asp Asp
        355                 360                 365

Leu Lys His Val Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp
    370                 375                 380

Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu
385                 390                 395                 400

Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly
                405                 410                 415

Arg Asp Phe Lys Arg Ala His His His His His His
            420                 425
```

<210> 9
<211> 1008
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1005)

<220>
<223> Description of Artificial Sequence: Nucleic acid encoding Hybrid of Group A Streptococci M protein and M-like protein

<400> 9

```
atg gtc gcg act agg tct cag aca gat act ctg gaa aaa gta caa gaa   48
Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu
  1               5                  10                  15

cgt gct gac aag ttt gag ata gaa aac aat acg tta aaa ctt aag aat   96
Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn
             20                  25                  30

agt gac tta agt ttt aat aat aaa gcg tta aaa gat cat aat gat gag   144
Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu
         35                  40                  45

tta act gaa gag ttg agt aat gct aaa gag aaa cta cgt cac gtg gcc   192
Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg His Val Ala
     50                  55                  60

gtg act agg ggt aca ata aat gac ccg caa aga gca aaa gaa gct ctt   240
Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu Ala Leu
 65                  70                  75                  80

gac aag tat gag cta gaa aac cat gac tta aaa act aag gga tcc aga   288
Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Gly Ser Arg
                 85                  90                  95

gtg ttt cct agg ggg acg gta gaa aac ccg gac aaa gca cga gaa ctt   336
Val Phe Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu
            100                 105                 110

ctt aac aag tat gac gta gag aac tct atg tta caa gct aat aat gac   384
Leu Asn Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp
        115                 120                 125

aag tta cca tgg aga gtg cgt tat act agg cat acg cca gaa gat aag   432
Lys Leu Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys
    130                 135                 140

cta aaa aaa att att gac gat ctt gac gca aaa gaa cat gaa tta caa   480
Leu Lys Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln
145                 150                 155                 160

caa cag aat gag aag tta tct ctg cag aaa gtg tat att act agg ggt   528
Gln Gln Asn Glu Lys Leu Ser Leu Gln Lys Val Tyr Ile Thr Arg Gly
                165                 170                 175
```

```
atg aca aaa gag gac gta gaa aaa att gct aac aac ctt gac ata gaa   576
Met Thr Lys Glu Asp Val Glu Lys Ile Ala Asn Asn Leu Asp Ile Glu
            180                 185                 190

aac cat ggg tta aaa caa cag aat gaa cag tta tct act gat aaa caa   624
Asn His Gly Leu Lys Gln Gln Asn Glu Gln Leu Ser Thr Asp Lys Gln
        195                 200                 205

ggt ctt gaa gaa cag aat ggt acc gat aga gtt agt agg tct atg tca   672
Gly Leu Glu Glu Gln Asn Gly Thr Asp Arg Val Ser Arg Ser Met Ser
    210                 215                 220

aga gat gat cta tta aac agg gct cag gat ctt gaa gca aaa aac cac   720
Arg Asp Asp Leu Leu Asn Arg Ala Gln Asp Leu Glu Ala Lys Asn His
225                 230                 235                 240

ggg tta gaa cac cag aat act aag tta tct act gaa aat aaa acg ctt   768
Gly Leu Glu His Gln Asn Thr Lys Leu Ser Thr Glu Asn Lys Thr Leu
                245                 250                 255

caa gaa caa gca gaa gca cgc cag aaa gaa atc gat gtc gcg act agg   816
Gln Glu Gln Ala Glu Ala Arg Gln Lys Glu Ile Asp Val Ala Thr Arg
            260                 265                 270

tct cag aca gat act ctg gaa aaa gta caa gaa cgt gct gac aag ttt   864
Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala Asp Lys Phe
        275                 280                 285

gag ata gaa aac aat acg tta aaa ctt aag aat agt gac tta agt ttt   912
Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser Phe
    290                 295                 300

aat aat aaa gcg tta aaa gat cat aat gat gag tta act gaa gag ttg   960
Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu
305                 310                 315                 320

agt aat gct aaa gag aaa cta cgt cac cac cac cac cac cac tgc tga  1008
Ser Asn Ala Lys Glu Lys Leu Arg His His His His His His Cys
                325                 330                 335
```

<210> 10
<211> 335
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 10

```
Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu
  1               5                  10                  15

Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn
             20                  25                  30
```

```
Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu
        35                  40                  45

Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg His Val Ala
    50                  55                  60

Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu Ala Leu
65                  70                  75                  80

Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Gly Ser Arg
                85                  90                  95

Val Phe Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu
            100                 105                 110

Leu Asn Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp
        115                 120                 125

Lys Leu Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys
    130                 135                 140

Leu Lys Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln
145                 150                 155                 160

Gln Gln Asn Glu Lys Leu Ser Leu Gln Lys Val Tyr Ile Thr Arg Gly
                165                 170                 175

Met Thr Lys Glu Asp Val Glu Lys Ile Ala Asn Asn Leu Asp Ile Glu
            180                 185                 190

Asn His Gly Leu Lys Gln Gln Asn Glu Gln Leu Ser Thr Asp Lys Gln
        195                 200                 205

Gly Leu Glu Glu Gln Asn Gly Thr Asp Arg Val Ser Arg Ser Met Ser
    210                 215                 220

Arg Asp Asp Leu Leu Asn Arg Ala Gln Asp Leu Glu Ala Lys Asn His
225                 230                 235                 240

Gly Leu Glu His Gln Asn Thr Lys Leu Ser Thr Glu Asn Lys Thr Leu
                245                 250                 255

Gln Glu Gln Ala Glu Ala Arg Gln Lys Glu Ile Asp Val Ala Thr Arg
            260                 265                 270

Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala Asp Lys Phe
        275                 280                 285

Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp Leu Ser Phe
    290                 295                 300

Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr Glu Glu Leu
305                 310                 315                 320

Ser Asn Ala Lys Glu Lys Leu Arg His His His His His His Cys
                325                 330                 335
```

<210> 11
<211> 1194
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1191)

<220>
<223> Description of Artificial Sequence: Hybrid of
Group A Streptococci M protein and M-like protein

<400> 11

```
atg agt aag aac cct gtc cct gtc aaa aaa gaa gca aaa tta agt gaa   48
Met Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu
  1               5                  10                  15

gca gaa tta cat gac aaa att aaa aac ctt gaa gag gaa aaa gca gaa   96
Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu
             20                  25                  30

tta ttc gag aaa tta gat aaa gtt gaa gaa gag cat aaa aaa gtt gaa  144
Leu Phe Glu Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Val Glu
         35                  40                  45

gaa gag cat gtc gac gaa gaa cac cct gac gtt gtc gct gct aga gaa  192
Glu Glu His Val Asp Glu Glu His Pro Asp Val Val Ala Ala Arg Glu
     50                  55                  60

agc gta cta aat aat gtc cgt gta ccg ggt aca ctt tgg cta cgt caa  240
Ser Val Leu Asn Asn Val Arg Val Pro Gly Thr Leu Trp Leu Arg Gln
 65                  70                  75                  80

aaa gaa gaa aat gac aaa ctt aaa ttg gaa aag aaa ggg ctt gag act  288
Lys Glu Glu Asn Asp Lys Leu Lys Leu Glu Lys Lys Gly Leu Glu Thr
                 85                  90                  95

gag tta cag gaa aag gaa caa gct agc gaa gaa gca tca aat aat ggg  336
Glu Leu Gln Glu Lys Glu Gln Ala Ser Glu Glu Ala Ser Asn Asn Gly
            100                 105                 110

caa ctc aca tta cag cat aaa aat aat gca ttg act agt gag aat gag  384
Gln Leu Thr Leu Gln His Lys Asn Asn Ala Leu Thr Ser Glu Asn Glu
        115                 120                 125

tct ctt aga aga gaa aaa gat aga tat ttg tat gaa aaa gaa gaa tta  432
Ser Leu Arg Arg Glu Lys Asp Arg Tyr Leu Tyr Glu Lys Glu Glu Leu
    130                 135                 140

gaa gga tcc gag tca tca aat aat gcg gag tca tca aac att tct caa  480
Glu Gly Ser Glu Ser Ser Asn Asn Ala Glu Ser Ser Asn Ile Ser Gln
145                 150                 155                 160

gaa agc aaa cta ata aat aca ttg act gat gaa aat gag aaa ctc aga  528
Glu Ser Lys Leu Ile Asn Thr Leu Thr Asp Glu Asn Glu Lys Leu Arg
                165                 170                 175
```

```
gaa gag ctc caa cag tat tat gca tta agt gat gct aaa gaa gaa gaa   576
Glu Glu Leu Gln Gln Tyr Tyr Ala Leu Ser Asp Ala Lys Glu Glu Glu
            180                 185                 190

cct agg tat aaa gca ctg cag act gaa gtt aag gct gcg ggg caa agc   624
Pro Arg Tyr Lys Ala Leu Gln Thr Glu Val Lys Ala Ala Gly Gln Ser
        195                 200                 205

gct cct aaa ggt aca aac gtg agc gca gac cta tat aat tcg cta tgg   672
Ala Pro Lys Gly Thr Asn Val Ser Ala Asp Leu Tyr Asn Ser Leu Trp
    210                 215                 220

gat gaa aat aaa act ctt aga gaa aaa caa gaa gag tat ata aca aaa   720
Asp Glu Asn Lys Thr Leu Arg Glu Lys Gln Glu Glu Tyr Ile Thr Lys
225                 230                 235                 240

att caa aat gaa gag aca aaa aat aaa ggt acc gaa caa gca aaa aat   768
Ile Gln Asn Glu Glu Thr Lys Asn Lys Gly Thr Glu Gln Ala Lys Asn
                245                 250                 255

aat aat ggg gaa ctc aca tta cag caa aaa tac gat gca ttg act aat   816
Asn Asn Gly Glu Leu Thr Leu Gln Gln Lys Tyr Asp Ala Leu Thr Asn
            260                 265                 270

gag aat aag tct ctt aga aga gag aga gat aac tat tta aat tat tta   864
Glu Asn Lys Ser Leu Arg Arg Glu Arg Asp Asn Tyr Leu Asn Tyr Leu
        275                 280                 285

tat gaa aaa cca tgg gaa gag cat gaa aaa gta aca caa gcc aga gaa   912
Tyr Glu Lys Pro Trp Glu Glu His Glu Lys Val Thr Gln Ala Arg Glu
    290                 295                 300

gcg gtt atc aga gag atg caa cag agg ggg aca aat ttt gga cct ctg   960
Ala Val Ile Arg Glu Met Gln Gln Arg Gly Thr Asn Phe Gly Pro Leu
305                 310                 315                 320

tta gca agt aca atg cga gat aat cac aat tta aaa gaa acg ctt gac  1008
Leu Ala Ser Thr Met Arg Asp Asn His Asn Leu Lys Glu Thr Leu Asp
                325                 330                 335

aaa act cac gtg agt aag aac cct gtc cct gtc aaa aaa gaa gca aaa  1056
Lys Thr His Val Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys
            340                 345                 350

tta agt gaa gca gaa tta cat gac aaa att aaa aac ctt gaa gag gaa  1104
Leu Ser Glu Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu
        355                 360                 365

aaa gca gaa tta ttc gag aaa tta gat aaa gtt gaa gaa gag cat aaa  1152
Lys Ala Glu Leu Phe Glu Lys Leu Asp Lys Val Glu Glu Glu His Lys
    370                 375                 380

aaa gtt gaa gaa gag cat cac cac cac cac cac cac tgc taa          1194
Lys Val Glu Glu Glu His His His His His His His Cys
385                 390                 395
```

<210> 12
<211> 397
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 12
Met Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys Leu Ser Glu
1 5 10 15

Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu Lys Ala Glu
20 25 30

Leu Phe Glu Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Val Glu
35 40 45

Glu Glu His Val Asp Glu Glu His Pro Asp Val Val Ala Ala Arg Glu
50 55 60

Ser Val Leu Asn Asn Val Arg Val Pro Gly Thr Leu Trp Leu Arg Gln
65 70 75 80

Lys Glu Glu Asn Asp Lys Leu Lys Leu Glu Lys Lys Gly Leu Glu Thr
85 90 95

Glu Leu Gln Glu Lys Glu Gln Ala Ser Glu Glu Ala Ser Asn Asn Gly
100 105 110

Gln Leu Thr Leu Gln His Lys Asn Asn Ala Leu Thr Ser Glu Asn Glu
115 120 125

Ser Leu Arg Arg Glu Lys Asp Arg Tyr Leu Tyr Glu Lys Glu Glu Leu
130 135 140

Glu Gly Ser Glu Ser Ser Asn Asn Ala Glu Ser Ser Asn Ile Ser Gln
145 150 155 160

Glu Ser Lys Leu Ile Asn Thr Leu Thr Asp Glu Asn Glu Lys Leu Arg
165 170 175

Glu Glu Leu Gln Gln Tyr Tyr Ala Leu Ser Asp Ala Lys Glu Glu Glu
180 185 190

Pro Arg Tyr Lys Ala Leu Gln Thr Glu Val Lys Ala Ala Gly Gln Ser
195 200 205

Ala Pro Lys Gly Thr Asn Val Ser Ala Asp Leu Tyr Asn Ser Leu Trp
210 215 220

Asp Glu Asn Lys Thr Leu Arg Glu Lys Gln Glu Glu Tyr Ile Thr Lys
225 230 235 240

Ile Gln Asn Glu Glu Thr Lys Asn Lys Gly Thr Glu Gln Ala Lys Asn
245 250 255

```
Asn Asn Gly Glu Leu Thr Leu Gln Gln Lys Tyr Asp Ala Leu Thr Asn
            260                 265                 270

Glu Asn Lys Ser Leu Arg Arg Glu Arg Asp Asn Tyr Leu Asn Tyr Leu
        275                 280                 285

Tyr Glu Lys Pro Trp Glu Glu His Glu Lys Val Thr Gln Ala Arg Glu
    290                 295                 300

Ala Val Ile Arg Glu Met Gln Gln Arg Gly Thr Asn Phe Gly Pro Leu
305                 310                 315                 320

Leu Ala Ser Thr Met Arg Asp Asn His Asn Leu Lys Glu Thr Leu Asp
                325                 330                 335

Lys Thr His Val Ser Lys Asn Pro Val Pro Val Lys Lys Glu Ala Lys
            340                 345                 350

Leu Ser Glu Ala Glu Leu His Asp Lys Ile Lys Asn Leu Glu Glu Glu
        355                 360                 365

Lys Ala Glu Leu Phe Glu Lys Leu Asp Lys Val Glu Glu Glu His Lys
    370                 375                 380

Lys Val Glu Glu Glu His His His His His His His His Cys
385                 390                 395
```

<210> 13
<211> 1212
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1209)

<220>
<223> Description of Artificial Sequence: Nucleic acid encoding Hybrid of Group A Streptococci M protein and M-like protein

<400> 13

```
atg agt gac aat att aat aga tct gtc tct gtc aaa gat aat gaa aaa   48
Met Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn Glu Lys
  1               5                  10                  15

gaa tta cat aac aaa att gca gac ctt gaa gag gaa agg ggt gaa cat   96
Glu Leu His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His
             20                  25                  30

cta gac aaa ata gat gaa cta aaa gaa gaa cta aaa gca aag gaa aaa   144
Leu Asp Lys Ile Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys
         35                  40                  45

agt tca gga tcc gct gat cac cct agc tat acc gct gct aaa gat gaa   192
Ser Ser Gly Ser Ala Asp His Pro Ser Tyr Thr Ala Ala Lys Asp Glu
```

```
     50                  55                  60

gta cta agt aag ttc tct gta ccg ggt cat gtt tgg gca cat gaa aga   240
Val Leu Ser Lys Phe Ser Val Pro Gly His Val Trp Ala His Glu Arg
 65                  70                  75                  80

gaa aaa aat gac aaa ctt agc tcg gaa aat gaa ggg ctt aag gct ggt   288
Glu Lys Asn Asp Lys Leu Ser Ser Glu Asn Glu Gly Leu Lys Ala Gly
                85                  90                  95

tta cag gaa aag gaa caa gct agc gaa ggg gtt tct gta ggt tca gat   336
Leu Gln Glu Lys Glu Gln Ala Ser Glu Gly Val Ser Val Gly Ser Asp
           100                 105                 110

gca tca cta cat aac cgc att aca gac ctt gaa gag gaa aga gaa aaa   384
Ala Ser Leu His Asn Arg Ile Thr Asp Leu Glu Glu Glu Arg Glu Lys
       115                 120                 125

tta tta aat aaa tta gat aaa gtt gaa gaa gag cat aaa aaa gat cat   432
Leu Leu Asn Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Asp His
   130                 135                 140

gaa caa ggt acc aac agt aag aac cct gcc cct gcc cct gcc tct gct   480
Glu Gln Gly Thr Asn Ser Lys Asn Pro Ala Pro Ala Pro Ala Ser Ala
145                 150                 155                 160

gtc cct gtc aaa aaa gaa gca aca aaa tta agt gaa gca gaa tta tat   528
Val Pro Val Lys Lys Glu Ala Thr Lys Leu Ser Glu Ala Glu Leu Tyr
               165                 170                 175

aac aaa att caa gaa ctt gaa gag gga aaa gca gaa tta ttc gga tcc   576
Asn Lys Ile Gln Glu Leu Glu Glu Gly Lys Ala Glu Leu Phe Gly Ser
           180                 185                 190

gaa gaa gaa cgt act ttt act gag tta cca tat gaa gca cga tac aaa   624
Glu Glu Glu Arg Thr Phe Thr Glu Leu Pro Tyr Glu Ala Arg Tyr Lys
       195                 200                 205

gca tgg aaa agt gaa aat gat gag ctt cgg gaa aat tat aga agg acc   672
Ala Trp Lys Ser Glu Asn Asp Glu Leu Arg Glu Asn Tyr Arg Arg Thr
   210                 215                 220

tta gat aag ttt aat act gag caa ggt aag act acg aga tta gaa gaa   720
Leu Asp Lys Phe Asn Thr Glu Gln Gly Lys Thr Thr Arg Leu Glu Glu
225                 230                 235                 240

caa aat aag ctt gcg gac gcg aac tcg aaa agc gtt tct aat agt aac   768
Gln Asn Lys Leu Ala Asp Ala Asn Ser Lys Ser Val Ser Asn Ser Asn
               245                 250                 255

gtg agc ata aat cta tat aat gag cta cag gct gaa cat gat aag cta   816
Val Ser Ile Asn Leu Tyr Asn Glu Leu Gln Ala Glu His Asp Lys Leu
           260                 265                 270

cag act aaa cat gag gag cta ttg gct gaa cat gat gct ctt aaa gaa   864
Gln Thr Lys His Glu Glu Leu Leu Ala Glu His Asp Ala Leu Lys Glu
       275                 280                 285
```

```
aaa caa gat aaa aat caa gaa ttc gat gac cgg agc gtt tct act aat   912
Lys Gln Asp Lys Asn Gln Glu Phe Asp Asp Arg Ser Val Ser Thr Asn
    290                 295                 300

agt ggt agc gtg agc aca cca tat aat aac cta ttg aat gaa tat gat   960
Ser Gly Ser Val Ser Thr Pro Tyr Asn Asn Leu Leu Asn Glu Tyr Asp
305                 310                 315                 320

gac cta ttg gct aaa cat ggt gag cta ttg agt gaa tat gat gct ctt   1008
Asp Leu Leu Ala Lys His Gly Glu Leu Leu Ser Glu Tyr Asp Ala Leu
                325                 330                 335

aaa gaa aaa caa gat aaa aat caa gaa gct agc agt gac aat att aat   1056
Lys Glu Lys Gln Asp Lys Asn Gln Glu Ala Ser Ser Asp Asn Ile Asn
            340                 345                 350

aga tct gtc tct gtc aaa gat aat gaa aaa gaa tta cat aac aaa att   1104
Arg Ser Val Ser Val Lys Asp Asn Glu Lys Glu Leu His Asn Lys Ile
        355                 360                 365

gca gac ctt gaa gag gaa agg ggt gaa cat cta gac aaa ata gat gaa   1152
Ala Asp Leu Glu Glu Glu Arg Gly Glu His Leu Asp Lys Ile Asp Glu
    370                 375                 380

cta aaa gaa gaa cta aaa gca aag gaa aaa agt tca cac cac cac cac   1200
Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser His His His His
385                 390                 395                 400

cac cac tgc tga                                                  1212
His His Cys
```

<210> 14
<211> 403
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 14

```
Met Ser Asp Asn Ile Asn Arg Ser Val Ser Val Lys Asp Asn Glu Lys
  1               5                  10                  15

Glu Leu His Asn Lys Ile Ala Asp Leu Glu Glu Glu Arg Gly Glu His
             20                  25                  30

Leu Asp Lys Ile Asp Glu Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys
         35                  40                  45

Ser Ser Gly Ser Ala Asp His Pro Ser Tyr Thr Ala Ala Lys Asp Glu
     50                  55                  60

Val Leu Ser Lys Phe Ser Val Pro Gly His Val Trp Ala His Glu Arg
 65                  70                  75                  80

Glu Lys Asn Asp Lys Leu Ser Ser Glu Asn Glu Gly Leu Lys Ala Gly
```

```
                85                  90                  95

Leu Gln Glu Lys Glu Gln Ala Ser Glu Gly Val Ser Val Gly Ser Asp
            100                 105                 110

Ala Ser Leu His Asn Arg Ile Thr Asp Leu Glu Glu Glu Arg Glu Lys
        115                 120                 125

Leu Leu Asn Lys Leu Asp Lys Val Glu Glu Glu His Lys Lys Asp His
    130                 135                 140

Glu Gln Gly Thr Asn Ser Lys Asn Pro Ala Pro Ala Pro Ala Ser Ala
145                 150                 155                 160

Val Pro Val Lys Lys Glu Ala Thr Lys Leu Ser Glu Ala Glu Leu Tyr
                165                 170                 175

Asn Lys Ile Gln Glu Leu Glu Glu Gly Lys Ala Glu Leu Phe Gly Ser
            180                 185                 190

Glu Glu Glu Arg Thr Phe Thr Glu Leu Pro Tyr Glu Ala Arg Tyr Lys
        195                 200                 205

Ala Trp Lys Ser Glu Asn Asp Glu Leu Arg Glu Asn Tyr Arg Arg Thr
    210                 215                 220

Leu Asp Lys Phe Asn Thr Glu Gln Gly Lys Thr Thr Arg Leu Glu Glu
225                 230                 235                 240

Gln Asn Lys Leu Ala Asp Ala Asn Ser Lys Ser Val Ser Asn Ser Asn
                245                 250                 255

Val Ser Ile Asn Leu Tyr Asn Glu Leu Gln Ala Glu His Asp Lys Leu
            260                 265                 270

Gln Thr Lys His Glu Glu Leu Leu Ala Glu His Asp Ala Leu Lys Glu
        275                 280                 285

Lys Gln Asp Lys Asn Gln Glu Phe Asp Asp Arg Ser Val Ser Thr Asn
    290                 295                 300

Ser Gly Ser Val Ser Thr Pro Tyr Asn Asn Leu Leu Asn Glu Tyr Asp
305                 310                 315                 320

Asp Leu Leu Ala Lys His Gly Glu Leu Leu Ser Glu Tyr Asp Ala Leu
                325                 330                 335

Lys Glu Lys Gln Asp Lys Asn Gln Glu Ala Ser Ser Asp Asn Ile Asn
            340                 345                 350

Arg Ser Val Ser Val Lys Asp Asn Glu Lys Glu Leu His Asn Lys Ile
        355                 360                 365

Ala Asp Leu Glu Glu Glu Arg Gly Glu His Leu Asp Lys Ile Asp Glu
    370                 375                 380

Leu Lys Glu Glu Leu Lys Ala Lys Glu Lys Ser Ser His His His His
385                 390                 395                 400
```

His His Cys

<210> 15
<211> 1290
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1287)

<220>
<223> Description of Artificial Sequence: Hybrid of
Group A Streptococci M protein and M-like protein

<400> 15

```
atg aac ggt gat ggt aat cct agg gaa gtt ata gaa gat ctt gca gca   48
Met Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala
  1               5                  10                  15

aac aat ccc gca ata caa aat ata cgt tta cgt cac gaa aac aag gac   96
Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp
             20                  25                  30

tta aaa gcg aga tta gag aat gca atg gaa gtt gca gga aga gat ttt   144
Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe
         35                  40                  45

aag aga gct ctc gac gat cat agt gat tta gtc gca gaa aaa caa cgt   192
Lys Arg Ala Leu Asp Asp His Ser Asp Leu Val Ala Glu Lys Gln Arg
     50                  55                  60

tta gaa gat tta gga caa aaa ttt gaa aga ctg aaa cag cgt tca gaa   240
Leu Glu Asp Leu Gly Gln Lys Phe Glu Arg Leu Lys Gln Arg Ser Glu
 65                  70                  75                  80

ctc tac ctt cag caa tac tat gat aat aaa tca aat gga tat aaa ggt   288
Leu Tyr Leu Gln Gln Tyr Tyr Asp Asn Lys Ser Asn Gly Tyr Lys Gly
                 85                  90                  95

gac tgg tat gta caa cag tta gga tcc gat tca gta agt gga tta gag   336
Asp Trp Tyr Val Gln Gln Leu Gly Ser Asp Ser Val Ser Gly Leu Glu
            100                 105                 110

gtg gca gac ccc tct gat agt aag aaa ctt att gaa tta ggt ttg gct   384
Val Ala Asp Pro Ser Asp Ser Lys Lys Leu Ile Glu Leu Gly Leu Ala
        115                 120                 125

aaa tac ctt aat gat aaa tta ccc ttt aaa act aaa gaa gat tca gag   432
Lys Tyr Leu Asn Asp Lys Leu Pro Phe Lys Thr Lys Glu Asp Ser Glu
    130                 135                 140

att tta tca gag tta cgt gat gta tta aaa aat ctg cag gag tct cca   480
Ile Leu Ser Glu Leu Arg Asp Val Leu Lys Asn Leu Gln Glu Ser Pro
```

```
145                 150                 155                 160

aaa agt act gag act tct gct aat gga gct gat aaa tta gct gat gca   528
Lys Ser Thr Glu Thr Ser Ala Asn Gly Ala Asp Lys Leu Ala Asp Ala
                165                 170                 175

tac aac aca ttg ctt act gaa cat gag aaa ctc aga gat gag tat tat   576
Tyr Asn Thr Leu Leu Thr Glu His Glu Lys Leu Arg Asp Glu Tyr Tyr
            180                 185                 190

aca tta att gat gct aaa gaa gaa gaa cct agg tat aaa gca ttg ggt   624
Thr Leu Ile Asp Ala Lys Glu Glu Glu Pro Arg Tyr Lys Ala Leu Gly
        195                 200                 205

acc ttg tta gat cag gtt aca caa tta tat act aaa cat aat agt aat   672
Thr Leu Leu Asp Gln Val Thr Gln Leu Tyr Thr Lys His Asn Ser Asn
    210                 215                 220

tac caa caa tat aat gca caa gct ggc aga ctt gac ctg aga caa aag   720
Tyr Gln Gln Tyr Asn Ala Gln Ala Gly Arg Leu Asp Leu Arg Gln Lys
225                 230                 235                 240

gct gaa tat cta aaa ggc ctt aat gat tgg gct gag agg ctg tta caa   768
Ala Glu Tyr Leu Lys Gly Leu Asn Asp Trp Ala Glu Arg Leu Leu Gln
                245                 250                 255

gag tta aat ggt acc aac aat gat ggt agg tct agg gac gtt acg gaa   816
Glu Leu Asn Gly Thr Asn Asn Asp Gly Arg Ser Arg Asp Val Thr Glu
            260                 265                 270

gag att gca gca aac aat acc aca gta caa aat ata cgt tta cgt aac   864
Glu Ile Ala Ala Asn Asn Thr Thr Val Gln Asn Ile Arg Leu Arg Asn
        275                 280                 285

gaa aac aag aac tta aaa gcg aaa aac gag gac tta gaa gcg aga tta   912
Glu Asn Lys Asn Leu Lys Ala Lys Asn Glu Asp Leu Glu Ala Arg Leu
    290                 295                 300

gag aat gca atg aat gtt gca gga aga gat ttt aag aga gct gaa ttc   960
Glu Asn Ala Met Asn Val Ala Gly Arg Asp Phe Lys Arg Ala Glu Phe
305                 310                 315                 320

gca cct ctt act cgt gct aca gca gac aat aaa gac gaa tta ata aaa  1008
Ala Pro Leu Thr Arg Ala Thr Ala Asp Asn Lys Asp Glu Leu Ile Lys
                325                 330                 335

aga gct aac ggt tat gag ata cag aac cat cag tta aca gtt gag aat  1056
Arg Ala Asn Gly Tyr Glu Ile Gln Asn His Gln Leu Thr Val Glu Asn
            340                 345                 350

aaa aaa tta aaa act gat aag gaa cag tta aca aaa gag aat gat gat  1104
Lys Lys Leu Lys Thr Asp Lys Glu Gln Leu Thr Lys Glu Asn Asp Asp
        355                 360                 365

tta aaa cac gtg aac ggt gat ggt aat cct agg gaa gtt ata gaa gat  1152
Leu Lys His Val Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp
    370                 375                 380
```

```
ctt gca gca aac aat ccc gca ata caa aat ata cgt tta cgt cac gaa   1200
Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu
385                 390                 395                 400

aac aag gac tta aaa gcg aga tta gag aat gca atg gaa gtt gca gga   1248
Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly
                405                 410                 415

aga gat ttt aag aga gct cac cac cac cac cac cac tgc taa           1290
Arg Asp Phe Lys Arg Ala His His His His His His Cys
            420                 425
```

<210> 16
<211> 429
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hybrid of Group A Streptococci M protein and M-like protein

<400> 16

```
Met Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala
  1               5                  10                  15

Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp
             20                  25                  30

Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe
         35                  40                  45

Lys Arg Ala Leu Asp Asp His Ser Asp Leu Val Ala Glu Lys Gln Arg
     50                  55                  60

Leu Glu Asp Leu Gly Gln Lys Phe Glu Arg Leu Lys Gln Arg Ser Glu
 65                  70                  75                  80

Leu Tyr Leu Gln Gln Tyr Tyr Asp Asn Lys Ser Asn Gly Tyr Lys Gly
                 85                  90                  95

Asp Trp Tyr Val Gln Gln Leu Gly Ser Asp Ser Val Ser Gly Leu Glu
            100                 105                 110

Val Ala Asp Pro Ser Asp Ser Lys Lys Leu Ile Glu Leu Gly Leu Ala
        115                 120                 125

Lys Tyr Leu Asn Asp Lys Leu Pro Phe Lys Thr Lys Glu Asp Ser Glu
    130                 135                 140

Ile Leu Ser Glu Leu Arg Asp Val Leu Lys Asn Leu Gln Glu Ser Pro
145                 150                 155                 160

Lys Ser Thr Glu Thr Ser Ala Asn Gly Ala Asp Lys Leu Ala Asp Ala
                165                 170                 175

Tyr Asn Thr Leu Leu Thr Glu His Glu Lys Leu Arg Asp Glu Tyr Tyr
            180                 185                 190
```

```
Thr Leu Ile Asp Ala Lys Glu Glu Glu Pro Arg Tyr Lys Ala Leu Gly
        195                 200                 205

Thr Leu Leu Asp Gln Val Thr Gln Leu Tyr Thr Lys His Asn Ser Asn
    210                 215                 220

Tyr Gln Gln Tyr Asn Ala Gln Ala Gly Arg Leu Asp Leu Arg Gln Lys
225                 230                 235                 240

Ala Glu Tyr Leu Lys Gly Leu Asn Asp Trp Ala Glu Arg Leu Leu Gln
                245                 250                 255

Glu Leu Asn Gly Thr Asn Asn Asp Gly Arg Ser Arg Asp Val Thr Glu
            260                 265                 270

Glu Ile Ala Ala Asn Asn Thr Thr Val Gln Asn Ile Arg Leu Arg Asn
        275                 280                 285

Glu Asn Lys Asn Leu Lys Ala Lys Asn Glu Asp Leu Glu Ala Arg Leu
    290                 295                 300

Glu Asn Ala Met Asn Val Ala Gly Arg Asp Phe Lys Arg Ala Glu Phe
305                 310                 315                 320

Ala Pro Leu Thr Arg Ala Thr Ala Asp Asn Lys Asp Glu Leu Ile Lys
                325                 330                 335

Arg Ala Asn Gly Tyr Glu Ile Gln Asn His Gln Leu Thr Val Glu Asn
            340                 345                 350

Lys Lys Leu Lys Thr Asp Lys Glu Gln Leu Thr Lys Glu Asn Asp Asp
        355                 360                 365

Leu Lys His Val Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp
    370                 375                 380

Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu
385                 390                 395                 400

Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly
                405                 410                 415

Arg Asp Phe Lys Arg Ala His His His His His His Cys
            420                 425


<210> 17
<211> 5347
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Vector

<400> 17
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
```

cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 120
ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 180
gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 360
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
tccgctcatg aattaattct tagaaaaact catcgagcat caaatgaaac tgcaatttat 600
tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat gaaggagaaa 660
actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc 720
gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta tcaagtgaga 780
aatcaccatg agtgacgact gaatccggtg agaatggcaa aagtttatgc atttctttcc 840
agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac 900
cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac 960
aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat 1020
tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgcag 1080
tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc ggaagaggca 1140
taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac 1200
ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg 1260
tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca 1320
tgttggaatt taatcgcggc ctagagcaag acgtttcccg ttgaatatgg ctcataacac 1380
cccttgtatt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa 1440
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 1500
gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg 1560
gtggtttgtt tgccggatca agagctacca actcttttte cgaaggtaac tggcttcagc 1620
agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 1680
aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc 1740
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 1800
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac 1860
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 1920
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt 1980
ccaggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 2040
cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat ggaaaaacgc cagcaacgcg 2100
gccttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta 2160
tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc 2220
agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg 2280
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatatggtgc actctcagta 2340
caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg 2400
ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct 2460
gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag 2520
gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc 2580
gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag 2640
aagcgttaat gtctggcttc tgataaagcg ggccatgtta agggcggttt tttcctgttt 2700
ggtcactgat gcctccgtgt aagggggatt tctgttcatg ggggtaatga taccgatgaa 2760
acgagagagg atgctcacga tacgggttac tgatgatgaa catgcccggt tactggaacg 2820
ttgtgagggt aaacaactgg cggtatggat gcggcgggac cagagaaaaa tcactcaggg 2880
tcaatgccag cgcttcgtta atacagatgt aggtgttcca cagggtagcc agcagcatcc 2940
tgcgatgcag atccggaaca taatggtgca gggcgctgac ttccgcgttt ccagacttta 3000
cgaaacacgg aaaccgaaga ccattcatgt tgttgctcag gtcgcagacg ttttgcagca 3060
gcagtcgctt cacgttcgct cgcgtatcgg tgattcattc tgctaaccag taaggcaacc 3120
ccgccagcct agccgggtcc tcaacgacag gagcacgatc atgcgcaccc gtggggccgc 3180
catgccggcg ataatggcct gcttctcgcc gaaacgtttg gtggcgggac cagtgacgaa 3240
ggcttgagcg agggcgtgca agattccgaa taccgcaagc gacaggccga tcatcgtcgc 3300
gctccagcga aagcggtcct cgccgaaaat gacccagagc gctgccggca cctgtcctac 3360
gagttgcatg ataaagaaga cagtcataag tgcggcgacg atagtcatgc cccgcgccca 3420
ccggaaggag ctgactgggt tgaaggctct caagggcatc ggtcgagatc ccggtgccta 3480
atgagtgagc taacttacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa 3540

cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat 3600
tgggcgccag ggtggttttt cttttcacca gtgagacggg caacagctga ttgcccttca 3660
ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc agcaggcgaa 3720
aatcctgttt gatggtggtt aacggcggga tataacatga gctgtcttcg gtatcgtcgt 3780
atcccactac cgagatatcc gcaccaacgc gcagcccgga ctcggtaatg gcgcgcattg 3840
cgcccagcgc catctgatcg ttggcaacca gcatcgcagt gggaacgatg ccctcattca 3900
gcatttgcat ggtttgttga aaaccggaca tggcactcca gtcgccttcc cgttccgcta 3960
tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc agacgcgccg 4020
agacagaact taatgggccc gctaacagcg cgatttgctg gtgacccaat gcgaccagat 4080
gctccacgcc cagtcgcgta ccgtcttcat gggagaaaat aatactgttg atgggtgtct 4140
ggtcagagac atcaagaaat aacgccggaa cattagtgca ggcagcttcc acagcaatgg 4200
catcctggtc atccagcgga tagttaatga tcagcccact gacgcgttgc gcgagaagat 4260
tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac accaccacgc 4320
tggcacccag ttgatcggcg cgagatttaa tcgccgcgac aatttgcgac ggcgcgtgca 4380
gggccagact ggaggtggca acgccaatca gcaacgactg tttgcccgcc agttgttgtg 4440
ccacgcggtt gggaatgtaa ttcagctccg ccatcgccgc ttccactttt tcccgcgttt 4500
tcgcagaaac gtggctggcc tggttcacca cgcgggaaac ggtctgataa gagacaccgg 4560
catactctgc gacatcgtat aacgttactg gtttcacatt caccaccctg aattgactct 4620
cttccgggcg ctatcatgcc ataccgcgaa aggttttgcg ccattcgatg gtgtccggga 4680
tctcgacgct ctcccttatg cgactcctgc attaggaagc agcccagtag taggttgagg 4740
ccgttgagca ccgccgccgc aaggaatggt gcatgcaagg agatggcgcc caacagtccc 4800
ccggccacgg ggcctgccac catacccacg ccgaaacaag cgctcatgag cccgaagtgg 4860
cgagcccgat cttccccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg 4920
gcgccggtga tgccggccac gatgcgtccg gcgtagagga tcgagatcta atcataaaaa 4980
atttatttgc tttgtgagcg gataacaatt ataatagatt caattgtgag cggataacaa 5040
ttataataga ttcaattcta aatttacaag aatttcacac agaattcatt aaagaggaga 5100
aattacatat ggctagcatg actggtggac agcaaatggg tcgcggatcc gaattcgagc 5160
tccgtcgaca agcttgcggc cgcactcgag caccaccacc accaccactg agatccggct 5220
gctaacaaag cccgaaagga agctgagttg gctgctgcca ccgctgagca ataactagca 5280
taaccccttg gggcctctaa acgggtcttg aggggttttt tgctgaaagg aggaactata 5340
tccggat 5347

<210> 18
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Carboxy terminal tag sequence

<400> 18
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 19
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Carboxy terminal tag sequence

<400> 19
Asp Leu Tyr Asp Asp Asp Asp Lys
1 5

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.